PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ :			(11	(11) International Publication Number: WO 98/42738				
C07K 14/00		A1	(43) International Publication Date: 1 October 1998 (01.10.98				
(21) International Appli	cation Number: PCT/US	598/0531	11	60/056,370 19 August 1997 (19.08.97) U 60/060,862 2 October 1997 (02.10.97) U				
(22) International Filing	Date: 19 March 1998	(19.03.98	8)					
(30) Priority Data: 60/041,281 60/041,276 60/042,344 60/041,277 60/048,355 60/048,096 60/048,154 60/048,160 60/048,186 60/048,187 60/048,187 60/048,099 60/050,937 60/048,135 60/048,352 60/048,135 60/048,352 60/048,350 60/048,350 60/054,804	21 March 1997 (21.03.97) 21 March 1997 (21.03.97) 21 March 1997 (21.03.97) 21 March 1997 (21.03.97) 30 May 1997 (30.05.97) 5 August 1997 (05.08.97)	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0		 (71) Applicant (for all designated States except US): HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key We Avenue, Rockville, MD 20850 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): YOUNG, Paul [US/US] 122 Beckwith Street, Gaithersburg, MD 20878 (US) GREENE, John, M. [US/US]; 872 Diamond Avenue Gaithersburg, MD 20878 (US). FERRIR, Ann, M. [US/US] 13203 L Astoria Hill Court, Germantown, MD 20874 (US) RUBEN, Steven, M. [US/US]; 18528 Heritage Hills Drive Olney, MD 20832 (US). ROSEN, Craig, A. [US/US] 22400 Rolling Hill Road, Laytonsville, MD 20882 (US) DUAN, Roxanne [US/US]; 4541 Fairfield Drive, Bethesd: MD 20814 (US). HU, Jing-Shan [CN/US]; 1247 Lakesid Drive #3034, Sunnyvale, CA 94086 (US). FLORENCE Kimberly, A. [US/US]; 12805 Atlantic Avenue, Rockville MD 20814 (US). OLSEN, Henrik, S. [DK/US]; 18 Kendrick Place #24, Gaithersburg, MD 20878 (US). EBNER, Reinhard [DE/US]; 9906 Shelbume Terrace #316 Gaithersburg, MD 20878 (US). BREWER, Laurie, A [US/US]; 14920 M. Nebo Road, Poolesville, MD 2083 (US). MOORE, Paul, A. [GB/US]; Apartment 104, 190 Holly Ridge Drive, McLean, VA 22102 (US). SHI, Yangg [CN/US]; 437 West Side Drive, Gaithersburg, MD 2087 (US). LAFLEUR, David, W. [US/US]; 1615 Q Street, N.W. #807, Washington, DC 20009 (US). NI, Jian [CN/US 5502 Manorfield Road, Rockville, MD 20853 (US). (74) Agents: BROOKES, Anders, A. et al.; Human Genom Sciences, Inc., 9410 Key West Avenue, Rockville, MI 10850 (US). (81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BH BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SI, TI, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIP patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasia patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), Europea patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, GR, IE, TI, UM, CR, LL, LY, LY, LY, LY, LY, LY, LY, LY, LY				
				Published With international search report.				

(54) Title: 87 HUMAN SECRETED PROTEINS

(57) Abstract

The present invention relates to 87 novel human secreted proteins and isolated nucleic acids containing the coding regions of the genes encoding such proteins. Also provided are vectors, host cells, antibodies, and recombinant methods for producing human secreted proteins. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating disorders related to these novel human secreted proteins.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GB	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Paso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
СН	Switzerland	KG	Kyrgyzstan	NO	Norway	zw	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
СМ	Cameroon		Republic of Korea	PL.	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
cz	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

10

15

20

25

30

35

87 Human Secreted Proteins

Field of the Invention

This invention relates to newly identified polynucleotides and the polypeptides encoded by these polynucleotides, uses of such polynucleotides and polypeptides, and their production.

Background of the Invention

Unlike bacterium, which exist as a single compartment surrounded by a membrane, human cells and other eucaryotes are subdivided by membranes into many functionally distinct compartments. Each membrane-bounded compartment, or organelle, contains different proteins essential for the function of the organelle. The cell uses "sorting signals," which are amino acid motifs located within the protein, to target proteins to particular cellular organelles.

One type of sorting signal, called a signal sequence, a signal peptide, or a leader sequence, directs a class of proteins to an organelle called the endoplasmic reticulum (ER). The ER separates the membrane-bounded proteins from all other types of proteins. Once localized to the ER, both groups of proteins can be further directed to another organelle called the Golgi apparatus. Here, the Golgi distributes the proteins to vesicles, including secretory vesicles, the cell membrane, lysosomes, and the other organelles.

Proteins targeted to the ER by a signal sequence can be released into the extracellular space as a secreted protein. For example, vesicles containing secreted proteins can fuse with the cell membrane and release their contents into the extracellular space - a process called exocytosis. Exocytosis can occur constitutively or after receipt of a triggering signal. In the latter case, the proteins are stored in secretory vesicles (or secretory granules) until exocytosis is triggered. Similarly, proteins residing on the cell membrane can also be secreted into the extracellular space by proteolytic cleavage of a "linker" holding the protein to the membrane.

Despite the great progress made in recent years, only a small number of genes encoding human secreted proteins have been identified. These secreted proteins include the commercially valuable human insulin, interferon, Factor VIII, human growth hormone, tissue plasminogen activator, and erythropoeitin. Thus, in light of the pervasive role of secreted proteins in human physiology, a need exists for identifying and characterizing novel human secreted proteins and the genes that encode them. This knowledge will allow one to detect, to treat, and to prevent medical disorders by using secreted proteins or the genes that encode them.

10

15

20

25

30

35

Summary of the Invention

The present invention relates to novel polynucleotides and the encoded polypeptides. Moreover, the present invention relates to vectors, host cells, antibodies, and recombinant methods for producing the polypeptides and polynucleotides. Also provided are diagnostic methods for detecting disorders related to the polypeptides, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying binding partners of the polypeptides.

Detailed Description

Definitions

The following definitions are provided to facilitate understanding of certain terms used throughout this specification.

In the present invention, "isolated" refers to material removed from its original environment (e.g., the natural environment if it is naturally occurring), and thus is altered "by the hand of man" from its natural state. For example, an isolated polynucleotide could be part of a vector or a composition of matter, or could be contained within a cell, and still be "isolated" because that vector, composition of matter, or particular cell is not the original environment of the polynucleotide.

In the present invention, a "secreted" protein refers to those proteins capable of being directed to the ER, secretory vesicles, or the extracellular space as a result of a signal sequence, as well as those proteins released into the extracellular space without necessarily containing a signal sequence. If the secreted protein is released into the extracellular space, the secreted protein can undergo extracellular processing to produce a "mature" protein. Release into the extracellular space can occur by many mechanisms, including exocytosis and proteolytic cleavage.

As used herein, a "polynucleotide" refers to a molecule having a nucleic acid sequence contained in SEQ ID NO:X or the cDNA contained within the clone deposited with the ATCC. For example, the polynucleotide can contain the nucleotide sequence of the full length cDNA sequence, including the 5' and 3' untranslated sequences, the coding region, with or without the signal sequence, the secreted protein coding region, as well as fragments, epitopes, domains, and variants of the nucleic acid sequence. Moreover, as used herein, a "polypeptide" refers to a molecule having the translated amino acid sequence generated from the polynucleotide as broadly defined.

In the present invention, the full length sequence identified as SEQ ID NO:X was often generated by overlapping sequences contained in multiple clones (contig

15

20

25

30

analysis). A representative clone containing all or most of the sequence for SEQ ID NO:X was deposited with the American Type Culture Collection ("ATCC"). As shown in Table 1, each clone is identified by a cDNA Clone ID (Identifier) and the ATCC Deposit Number. The ATCC is located at 10801 University Boulevard, Manassas, Virginia 20110-2209, USA. The ATCC deposit was made pursuant to the terms of the Budapest Treaty on the international recognition of the deposit of microorganisms for purposes of patent procedure.

A "polynucleotide" of the present invention also includes those polynucleotides capable of hybridizing, under stringent hybridization conditions, to sequences contained in SEQ ID NO:X, the complement thereof, or the cDNA within the clone deposited with the ATCC. "Stringent hybridization conditions" refers to an overnight incubation at 42° C in a solution comprising 50% formamide, 5x SSC (750 mM NaCl, 75 mM sodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and $20~\mu g/ml$ denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65° C.

Also contemplated are nucleic acid molecules that hybridize to the polynucleotides of the present invention at lower stringency hybridization conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency); salt conditions, or temperature. For example, lower stringency conditions include an overnight incubation at 37°C in a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 0.2M NaH₂PO₄; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 ug/ml salmon sperm blocking DNA; followed by washes at 50°C with 1XSSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5X SSC).

Note that variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include Denhardt's reagent, BLOTTO, heparin, denatured salmon sperm DNA, and commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

Of course, a polynucleotide which hybridizes only to polyA+ sequences (such as any 3' terminal polyA+ tract of a cDNA shown in the sequence listing), or to a

10

15

20

25

30

35

complementary stretch of T (or U) residues, would not be included in the definition of "polynucleotide," since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone).

The polynucleotide of the present invention can be composed of any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of single-and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, the polynucleotide can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. A polynucleotide may also contain one or more modified bases or DNA or RNA backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications can be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

The polypeptide of the present invention can be composed of amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched, for example, as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine,

20

30

35

formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. (See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993); POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, pgs. 1-12 (1983); Seifter et al., Meth Enzymol 182:626-646 (1990); Rattan et al., Ann NY Acad Sci 663:48-62 (1992).)

"SEQ ID NO:X" refers to a polynucleotide sequence while "SEQ ID NO:Y" refers to a polypeptide sequence, both sequences identified by an integer specified in Table 1.

"A polypeptide having biological activity" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. In the case where dose dependency does exist, it need not be identical to that of the polypeptide, but rather substantially similar to the dose-dependence in a given activity as compared to the polypeptide of the present invention (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about tenfold less activity, and most preferably, not more than about three-fold less activity relative to the polypeptide of the present invention.)

25 Polynucleotides and Polypeptides of the Invention

FEATURES OF PROTEIN ENCODED BY GENE NO: 1

The translation product of this gene shares sequence homology with nucleolin, which is thought to be important in macromolecule binding, as well as some membrane proteins. Preferred polypeptide fragments comprise the amino acid sequence:

DPEAADSGEPQNKRTPDLPEEEYVKEEIQENEEAVKKMLVEATREFEEVVVDES
(SEQ ID NO:239); QKLKRKAEEDPEAADSGEPQNKRTPDLPEEEYVKEEIQENEE
AVKKMLVEATREFEEVVVDES (SEQ ID NO:240); KAMEKSSLTQHSWQSLKDR
YLKHLRGQEHKYLLGDAPVSPSSQKLKRKAEEDPEAADSGEPQNKRTPDLPEE
EYVKEEIQENEEAVKKMLVEATREFEEVVVDESPPDFEIHI (SEQ ID NO:241).
Also preferred are the polynucleotide fragments encoding these polypeptide fragments.

10

15

20

25

30

35

This gene maps to chromosome 16, and therefore can be used as a marker in linkage analysis for chromosome 16.

This gene is expressed primarily in brain and kidney and to a lesser extent in wide range of tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, cell-cell interaction or cell-matrix interaction. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain and kidney, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., brain and other tissue of the nervous system, and kidney, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:125 as residues: Met-1 to Trp-10.

The tissue distribution and homology to nucleolin indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment/diagnosis of diseases involving cell-cell interaction or cell-extracellular matrix interaction.

FEATURES OF PROTEIN ENCODED BY GENE NO: 2

The translation product of this gene shares sequence homology with a porcine zona pellucida protein ZPDS.1711. (See Accession No. R39356.) These two proteins have weak homology with *Drosophila* commissureless and metal homeostasis proteins which are thought to be important in controlling growth cone guidance across the CNS midline and protecting cells against reactive oxygen toxicity. thus, based on homology, it is likely that this gene also be involved in development. Preferred polypeptide fragments comprise the amino acid sequence: LPSYDEAERTKAEATIPLVPGRDEDF VGRDDFDDADQLRIGNDGIFMLTFFMAFLFNWIGFFLSFCLTTSAAGRYGAISG FGLSLIKWILIVRFSTYFPGYFDGQYWLWWVFLVLGFLLFLRGFINYAKVRKM PETFSNLPRTRVLFI (SEQ ID NO:242); and/or AGRYGAISGFGLSLIKWILIVRFS (SEQ ID NO:243). Also preferred are polynucleotide fragments encoding these polypeptide fragments. This gene maps to chromosome 5, and therefore can be used in linkage analysis as a marker for chromosome 5.

10

15

20

30

35

This gene is expressed primarily in kidney, adrenal gland, brain and to a lesser extent in wide range of tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, fertilization control or tissue damages by metabolites or other toxic agents. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive and urosecretion system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., kidney, adrenal gland, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to zona pellucida protein indicates that polynucleotides and polypeptides corresponding to this gene are useful for fertility control such as controceptive development. The homology with metal homeostasis and commissureless genes indicates the gene's function in spermatozoa guidance and protection. It would also be useful for the treatment/diagnosis of tissue damages caused by toxic metabolites and other agents since the gene product is also expressed in urosecretive tissues.

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 3

This gene is expressed primarily in liver and to a lesser extent in placenta. Preferred polypeptide fragments comprise the amino acid sequence: MKHLSAWNFT KLTFLQLWEI FEGSVENCQTLTSYSKLQIKYTFSRGSTFYI (SEQ ID NO:244). Also preferred are polynucleotide fragments encoding these polypeptide fragments.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, digestive and nutrient transport/utilization disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the digestive and

10

35

circulatory system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., liver, and placenta, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in liver and placenta indicates that the protein product is either an extracellular enzyme or a molecule carrier. Therefore, polynucleotides and polypeptides corresponding to this gene are useful for diagnosis/treatment of digestive and nutrient transport/utilization disorders, including malabsorption and malnutrition.

FEATURES OF PROTEIN ENCODED BY GENE NO: 4

This gene shares homology with the sap47 gene of Drosophila melanogaster, a gene which codes for a conserved neuronal protein associated with synaptic terminals. (See 15 Mol. Brain Res. 32:45-54 (1995); see also, Accession No. 929571.) Thus, based on homology, the gene of the present invention also should be associated with synaptic terminals. Preferred polypeptide fragments comprise the amino acid sequence: FSSDFRTSPWESRRVESKATSARCGLWGSGPRRRPASGMFRGLSSWLGLQOP VAGGGQPNGDAPPEQPSETVAESAEEELQQAGDQELLHQAKDFGNYLFNFASA 20 ATKKITESVAETAQTIKKSVEEGKIDGIIDKTIIGDFQKEQKKFVEEQHTKKSEA AVPPWVDTNDEETIQQQILALSADKRNFLRDPPAGVQFNFDFDQMYPVALVML (SEQ ID NO:245); MRFALVPKLVKEEVFWRNYFYRVSLIKQSAQLTALAAQQQA AGKGGEEQ (SEQ ID NO:246); STSPGVSEFVSDAFDACNLNOEDLRKEMEOL VLDKKQEETAVLEEDSADWEKELQQELQEYEVVTESEKRDENWDK (SEQ ID 25 NO:247); SPWESRRVESKATSARCGLWGSGPRRRPASGMFRGLSSWLGLQQ PVAGGGQPNGDAPPEQPS (SEQ ID NO:248); PVAGGGQPNGDAPPEQPSETV ESAEELQQAGDQELLHQAKDFGNYLFNFASAATKKITESVAE (SEQ ID NO: 249); and/or FQKEQKKFVEEQHTKKSEAAVPPWVDTNDEETIQQQILALSADKR NFLRDPPAGVQFNFDFDQMYPVALVML (SEQ ID NO:250). Also preferred are 30 polynucleotide fragments encoding these polypeptide fragments.

This gene is expressed primarily in kidney pyramids and to a lesser extent in lung and other tissues of various types. This gene fluxes calcium in human aortic smooth muscle cells, and therefore is involved in signal transduction.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are

10

15

20

25

30

35

not limited to, renal and nervous disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the kidney and/or nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., kidney, lung, brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in kidney and lung and homology with sap47 indicates that the protein product has regulatory or direct functions in molecular exchange with body fluids and nervous system signaling. Polynucleotides and polypeptides corresponding to this gene are useful for treatment of disorders in kidney and nervous system.

FEATURES OF PROTEIN ENCODED BY GENE NO: 5

The translation product of this gene shares sequence homology with the mouse Ly-9.2 antigen which is thought to be an important cell surface marker in lymphoids, myeloids and hematopoietic progenitors. (See Accession No. gil198932.) Preferred polypeptide fragments comprise the amino acid sequence: PFICVARNPVSRNFSSPI LARKLCEGAA (SEQ ID NO:251); and/or KEDPANTVYSTVEIPKKMENPHSLLT MPDTPRL (SEQ ID NO:252). Also preferred are polynucleotide fragments encoding these polypeptide fragments. Based on homology, it is likely that this gene is also a cell surface marker, involved in hematopoiesis.

This gene is expressed primarily in activated macrophages, monocytes and T-cells and to a lesser extent in spleen and bone marrow.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune and hematopoietic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hematopoietic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., blood cells, and bone marrow, and cancerous and wounded

10

15

20

25

30

35

tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:129 as residues: Lys-26 to Tyr-33, Arg-44 to Ile-49, Ser-53 to Lys-71, Lys-86 to Pro-91.

The tissue distribution and homology to Ly-9.2 surface immunoglobulin family indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis of immune and hematopoietic disorders. Polypeptides and polynucleotides corresponding to this gene are also be used as a marker for leukemia or a modulator of the functions of the cells of macrophage/monocyte or T-cell types.

FEATURES OF PROTEIN ENCODED BY GENE NO: 6

The translation product of this gene shares sequence homology with the *Drosophila* glutactin gene which is thought to be important in cell-cell interaction or cell-extracellular matrix contact.

This gene is expressed primarily in colon tissue, aorta endothelial cells and to a lesser extent in skin, breast tissue and T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of these tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, diseases of the gastrointestinal tract, vascular system or T-cell development. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of these tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the digestive system, cardiovascular system, and immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., colon, cardiovascular tissue, skin, mammary tissue, and blood cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to glutactin indicates that polynucleotides and polypeptides corresponding to this gene are useful for the development and maintenance of the integrity of the basal membrane in the gastrointestinal tract and

10

15

20

25

30

35

cardiovascular system. The expression in T-cells also indicate the protein may be involved in T-cell adhesion, cell-cell interaction and development.

FEATURES OF PROTEIN ENCODED BY GENE NO: 7

The translation product of this gene shares sequence homology with MURF4 protein, an ATPase homolog, which is thought to be important in ATP hydrolysis.

This gene is expressed primarily in breast tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, breast cancer and non-neoplastic breast diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of these tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the breast tissue, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., mammary tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to MURF4 gene indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of neoplastic or non-neoplastic breast diseases because ATPase like protein may be involved in changed metabolic states of the breast.

FEATURES OF PROTEIN ENCODED BY GENE NO: 8

This gene shares homology to the alcohol dehydrogenase gene. Preferred polypeptide fragments comprise comprise the amino acid sequence: ASAVLLDLPNSG GEAQAKKLGNNCVFAPADVTSEKDVQTALALAKGKFGRVDVAVNCAGIAVAS KTYNLKKGQTHTLEDFQRVLDVNLMGTFNVIRLVAGEMGQNEPDQGGQRGVI INTASVAAFEGQVGQAAYSASKGGIVGMTLPIARDLAPIGIRVMTIAPGLFGTPL LTSLPEKVCNFLASQVPFPSRLGDPAEYAHLVQAIIENPFLNGEVIRLDGAIRMQ P (SEQ ID NO:253); and/or SVAAFEGQVGQAAYSASKGGIVGMTLPIA (SEQ ID NO:254). Polynucleotides encoding these fragements are also encompassed by the invention. Other groups have also recently cloned this gene, recognizing its homology to alcohol dehydrogenase. (See Accession No. 1778355.) Moreover, a second group

recently cloned the mouse homologue of this gene. (See Accession No. 2078284.) They found that the mouse homologue binds to amyloid beta-peptide and mediates neurotoxicity in Alzheimer's disease, calling the protein ERAB. This gene maps to chromosome X, and therefore can be used in linkage analysis as a marker for chromosome X. Therefore, mutations in the translated product of this gene may be involved in Alzheimer's disease in humans, as well as other sex linked diseases. This gene can be used as a diagnostic marker for these diseases.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:132 as residues: Arg-45 to Ser-53.

10

15

20

25

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 9

The translation product of this gene shares week sequence homology with rat N-methyl-D-aspartate receptor subunit and other proline-rich proteins which are thought to be important in neurotransmission or protein-protein intereaction.

This gene is expressed primarily in synovial hypoxia and to a lesser extent in ovary, senescent cells and brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, synovial hypoxia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the synovia and brain, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., synovial tissue, ovary and other reproductive tissue, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in synovial hypoxia and nerve tissues, and homology to N-methyl-D-aspartate receptor subunit and other proline-rich proteins indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and intervention of synovial hypoxia and other synovial disorders.

35

30

10

15

20

25

30

35

FEATURES OF PROTEIN ENCODED BY GENE NO: 10

This gene is expressed primarily in prostate and to a lesser extent in placenta and ovary.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, male and female infertility, cancer, and other hyperproliferative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of these tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system and neoplasia, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., prostate, placenta, ovary and other reproductive tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:134 as residues: Pro-17 to Met-23, Ala-30 to Trp-38, Ile-49 to Trp-54, Lys-68 to Gly-74, Thr-93 to Gly-99, Met-126 to Glu-132, Gly-173 to Ser-178, Lys-205 to Tyr-214.

The tissue distribution of this gene in the prostate, placenta and ovary indicates that this gene product is useful for treatment/diagnosis of male or female infertility, endocrine disorders, fetal deficiencies, ovarian failure, amenorrhea, ovarian cancer, benign prostate hyperplasia, prostate cancer, and other forms of cancer of the reproductive system.

FEATURES OF PROTEIN ENCODED BY GENE NO: 11

This gene is expressed primarily in the thyroid and to a lesser extent in the pineal gland. This gene maps to chromosome 10, and therefore can be used as a marker in linkage analysis for chromosome 10. Preferred polypeptide fragments comprise the amino acid sequence: HPIEWAINAATLSQFY (SEQ ID NO:256); CWIKYCLTLMQN AQLSMQDNIG (SEQ ID NO:257); KVSYLRPLDFEEARELFLLGQHYVF (SEQ ID NO:258); MERRCKMHKRXIAMLEPLTVDLNPQ (SEQ ID NO:259); and/or SHIV KKINNLNKSALKY YQLFLD (SEQ ID NO:260). Also preferred are polynucleotides encoding these polypeptide fragments.

Therefore, polynucleotides and polypeptides of the invention are useful as

10

15

20

DAIGNOCID JAIN 0040700A4

reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune, thyroid and pineal gland disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of these tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and endocrine systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., thyroid and pineal gland, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:135 as residues: Ser-2 to Ser-8, Thr-38 to Arg-44.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treating/detecting immune disorders such as arthritis, asthma, immune deficiency diseases (e.g., AIDS), and leukemia, as well as treating/detecting thymus disorders (e.g., Graves Disease, lymphocytic thyroiditis, hyperthyroidism, and hypothyroidism), and treating/detecting pineal gland disorders (e.g., circadian rhythm disturbances associated with shift work, jet lag, blindness, insomnia and old age).

FEATURES OF PROTEIN ENCODED BY GENE NO: 12

This gene is expressed primarily in lung and tonsils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, pulmonary or immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of these tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the pulmonary and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., pulmonary tissue, and tonsils, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily

10

15

20

25

30

35

fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:136 as residues: Glu-28 to Gly-49.

The tissue distribution of this gene only in lung indicates that it could play a role in the treatment/detection of lung lymphoma or sarcoma formation, pulmonary edema and embolism, bronchitis and cystic fibrosis. Its expression in tonsils indicates a potential role in the treatment/detection of immune disorders such as arthritis, asthma, immune deficiency diseases (e.g., AIDS), and leukemia, in addition to the treatment/detection of tonsillitis.

FEATURES OF PROTEIN ENCODED BY GENE NO: 13

This gene is expressed primarily in lymphoid, myeloid and erythroid cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, hematopoietic and immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of these tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoietic and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, myeloid cells, and bone marrow, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The predominant tissue distribution of this gene in hematopoietic cell types indicates that the gene could be important for the treatment or detection of immune or hematopoietic disorders including arthritis, asthma, immunodeficiency diseases and leukemia. Preferred embodiments of the present invention are polypeptide fragments comprising the amino acid sequence: FTHLSTCLLSLLLVRMSGFLLLARASPSI CALDSSCFVEYCSSYSSSCFLHQHFPSLLDHLCQ (SEQ ID NO:261); or FLLL ARASPSICALDSSCFVQEY (SEQ ID NO:262). Also preferred are polynucleotide fragments encoding these polypeptide fragments.

FEATURES OF PROTEIN ENCODED BY GENE NO: 14

This gene is homologous to the *Drosophila Regena* (Rga) gene. (See Accession No. 1658504.) This *Drosophila* gene is thought to be a homolog of the global negative

10

15

20

25

35

transcriptional regulator NOT2 (CDC36) from yeast, which modifies gene expression and suppresses position effect variegation. Preferred polypeptide fragments comprise the amino acid sequence: PDGRVTNIPQGMVTDQFGMIGLLTFIRAAETDPGMVHL ALGSDLTTLGLNLNS (SEQ ID NO:263); VHLALGSDLTTLGLNLNSPENLYP (SEQ ID NO:265); EDLLFYLYYMNGGDVLQLLAAVELFNRDWRYHKEERVWI TR (SEQ ID NO:264); and/or HNEDFPALPGS (SEQ ID NO:266).

This gene is expressed primarily in placenta and to a lesser extent in infant brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, neurodegenerative and developmental disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the neurological system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., placenta, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:138 as residues: Leu-9 to Tyr-15, Asp-34 to Gln-46, Pro-51 to Asp-57, Gly-88 to Thr-104, Thr-123 to Ser-128.

The tissue distribution of this gene indicates that it could be used in the detection and/or treatment of neurological disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, and panic disorder.

30 FEATURES OF PROTEIN ENCODED BY GENE NO: 15

This gene is expressed primarily in adrenal gland tumor and osteoclastoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, endocrine and bone disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for

10

15

20

25

30

35

differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine system and in bone, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., adrenal gland, and bone, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:139 as residues: Ile-52 to Trp-57.

The tissue distribution of this gene indicates that it may be involved in the treatment and/or detection of adrenal gland tumors, osteosarcomas, endocrine disorders and bone disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 16

The translation product of this gene shares sequence homology with the FK506 binding protein, a protein which plays an important role in immunosupression. (See Accession No. M75099.) Specifically, a 12-kDa FK506-binding protein (FKBP-12) is a cytosolic receptor for the immunosuppressants FK506 and rapamycin. (See, Proc. Natl. Acad. Sci. 88: 6677-6681 (1991).) Thus, based on homology, it is likely that this gene also has immunosuppression activity. Preferred polypeptides comprise the amino acid sequence: GRIIDTSLTRDPLVIELGQKQVIPGLEQSLLDMCVGEKRRAIIPSH LAYGKRGFPPSVPADAVVQYDVELIALIR (SEQ ID NO:267); and/or IHYTGSLV DGR IIDTS (SEQ ID NO:268). Also preferred are the polynucleotide fragments encoding these polypeptides.

This gene is expressed primarily in melanocytes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, cancer and other hyperproliferative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and cancer, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., melanocytes, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to

10

15

20

25

30

35

the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:140 as residues: Ala-118 to Phe-124, Arg-178 to Lys-201.

The tissue distribution and homology to the FK506 binding proteins which are believed to a role in immunosupression mediated by the immunosupressant drugs rapamycin and cyclosporin, indicates that this gene could serve as a novel target for the identification of novel immunosupressant drugs.

FEATURES OF PROTEIN ENCODED BY GENE NO: 17

The translation product of this gene shares sequence homology with the rat calcium-activated potassium channel rSK3, which is thought to be important in regulating vascular tone. (See Accession No. gil2564072, gil1575663, and gil1575661.) Although homologous to these proteins, this gene contains an 18 amino acid insert, not previously identified in the homologs. Preferred polypeptide fragments comprise the amino acid sequence: CESPESPAQPSGSSLPAWYH (SEQ ID NO:269). Also preferred are the polynucleotide fragments encoding these polypeptides.

This gene is expressed primarily in B-cells, frontal cortex and endothelial cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, cardiovascular (hyper/hypotension, asthma, pulmonary edema, pneumonia, heart disease, restenosis, atherosclerosis, stoke, angina and thrombosis) and neurological disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cardiovascular and nervous systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, brain and other tissue of the nervous system, and endothelium, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:141 as residues: Glu-72 to Gly-82, His-90 to Val-95, Gln-168 to Lys-174, Val-202 to Ser-212.

10

15

20

25

30

The tissue distribution and homology to calcium-activated potassium channels indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of vascular disorders (hyper/hypotension, athesma, pulmonary edema, pneumonia, heart disease, restenosis, atherosclerosis, stoke, angina and thrombosis).

FEATURES OF PROTEIN ENCODED BY GENE NO: 18

This gene is expressed primarily in smooth muscle and to a lesser extent in brain (amygdala, corpus colosum, hippocampus).

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, cardiovascular (hypertension, heart disease, athesma, pulmonary edema, -restenosis, atherosclerosis, stoke, angina, thrombosis, and wound healing), and neurological disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cardiovascular and neurological systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., smooth muscle, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:142 as residues: Lys-43 to Arg-49, Tyr-58 to Glu-65.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of cadiovascular disorders (hypertension, heart disease, athesma, pulmonary edema, restenosis, atherosclerosis, stoke, angina, thrombosis, and wound healing). Expression in brain indicates a role in the treatment and diagnosis of behavioral or neurological disorders, such as depression, schizophrenia, Alzheimer's disease, mania, dementia, paranoia, and addictive behavior.

35 FEATURES OF PROTEIN ENCODED BY GENE NO: 19

This gene is expressed primarily in T-cells (Jurkats, resting, activated, and

10

15

20

25

30

35

anergic T-cells), endothelial cells, pineal gland, and to a lesser extent in a variety of other tissues and cell types. Preferred polypeptide fragments comprise the amino acid sequence: EEAGAGRRCSHGGARPAGLGNEGLGLGGDPDHTDTGSRSKQRINN WKESKHKVIMASASARGNQDKDAHFPPPSKQSLLFCPKSKLHIHRAEISK (SEQ ID NO:270); and/or SKQRINNWKESKHKVIMASASAR (SEQ ID NO:271). Also preferred are the polypucleotide fragments encoding these polypepides.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, inflammation, immune and cardiovascular disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of these tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, neurological and vascular systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., T-cells and other blood cells, endothelial cells, and pineal gland, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:143 as residues: Phe-71 to Arg-76, Pro-82 to His-87, Glu-103 to Ala-111.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of immune disorders including: leukemias, lymphomas, auto-immune, immuno-supressive (e.g. transplantation) and immunodeficiencies (e.g. AIDS) and hematopoietic disorders. In addition, expression in the pineal gland might suggest a role in the diagnosis of specific brain tumors and treatment of neurological disorders. Endothelial cell expression might suggest a role in cadiovascular or respiratory/pulmonary disorders or infections (athesma, pulmonary edema, pneumonia).

FEATURES OF PROTEIN ENCODED BY GENE NO: 20

This gene is expressed primarily in brain and embryo and to a lesser extent in leukocytes. This gene maps to chromosome 15, and therefore can be used as a marker in linkage analysis to chromosome 15.

Therefore, polynucleotides and polypeptides of the invention are useful as

10

20

25

30

35

reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, developmental and neurological disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:144 as residues: Met-1 to Gly-8.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of immune 15 disorders including: leukemias, lymphomas, auto-immune, immuno-supressive (e.g. transplantation) and immunodeficiencies (e.g. AIDS) and hematopoietic disorders. The expression in the brain -- and in particular the fetal brain -- would suggest a possible role in the treatment and diagnosis of developmental and neurodegenerative diseases of the brain and nervous system (depression, schizophrenia, Alzheimer's disease, mania, dementia, paranoia, and addictive behavior).

FEATURES OF PROTEIN ENCODED BY GENE NO: 21

This gene is expressed primarily in brain, kidney, lung, liver, spleen, and a variety of leukocytes (especially T-cells) and to a lesser extent in a variety of other tissues and cell types.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, leukemias, lymphomas, autoimmune, immunosuppressive, and immunodeficiencies, hematopoietic disorders, as well as renal disorders, and neoplasms. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the renal, pulmonary, immune, and central nervous systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.,

10

15

20

25

30

35

brain and other tissue of the nervous system, kidney, pulmonary tissue, liver, spleen, and blood cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of renal conditions, such as acture renal failure, kidney fibrosis, and kidney tubule regeneration. The expression in leukocytes and other immune tissues indicates a role in immune disorders including: leukemias, lymphomas, auto-immune, immuno-supressive (e.g. transplantation) and immunodeficiencies (e.g. AIDS) and hematopoietic disorders. The expression in the brain -- and in particular the fetal brain -- indicates a possible role in the treatment and diagnosis of developmental and neurodegenerative diseases of the brain and nervous system (depression, schizophrenia, Alzheimer's disease, mania, dementia, paranoia, and addictive behavior).

FEATURES OF PROTEIN ENCODED BY GENE NO: 22

This gene is expressed primarily in skin (fetal epithelium, keratinocytes and skin). This gene also maps to chromosome 19, and therefore can be used in linkage analysis as a marker for chromosome 19.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, skin cancers (e.g., melanomas), eczema, psoriasis or other disorders of the skin. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of these tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skin, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., skin and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:146 as residues: Pro-28 to Glu-35, Ser-39 to Phe-44, Ala-94 to Gln-99.

15

20

25

30

35

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of skin cancers (e.g., melanomas), eczema, psoriasis or other disorders of the skin.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 23

This gene maps to chromosome 11. Another group recently isolated this same gene, associating the sequence to the region thought to harbor the gene involved in Multiple Endocrine Neoplasia Type 1, or MEN 1. (See Accession No. 2529721 and Genome Res. 7(7), 725-735 (1997), incorporated herein by reference in its entirety.) Preferred polypeptide fragments comprise the amino acid sequence: LFHWACLNERA AQLPRNTAXAGYQCPSCNGPS (SEQ ID NO:272).

This gene is expressed primarily in epididymus, pineal gland, T-cells, as well as fetal epithelium, lung and kidney.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune, metabolic mediated disorders, and MEN. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, renal, neurological and pulmonary systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., epididymus and other reproductive tissue, pineal gland, T-cells and other blood cells, epithelium, lung, and kidney, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of developmental deficiencies or abnormalities as well as a host of different disorders which arise as a result of conditions in the indicated tissues or cell types. An area of particular interest is in the treatment and diagnosis of immune disorders including: leukemias, lymphomas, auto-immune, immuno-supressive (e.g. transplantation) and immunodeficiencies (e.g. AIDS) and hematopoietic disorders. The expression in the brain, and in particular the fetal brain, would suggest a possible role in the treatment and diagnosis of

10

15

20

25

30

35

developmental and neurodegenerative diseases of the brain and nervous system (depression, schizophrenia, Alzheimer's disease, mania, dementia, paranoia, and addictive behavior). Respiratory/pulmonary disorders, such as athesma, pulmonary edema are also potential therapeutic areas, as well as renal conditions such as acute renal failure, kidney fibrosis and kidney tubule regeneration. Moreover, this gene can be used in the treatment and/or detection of MEN I.

FEATURES OF PROTEIN ENCODED BY GENE NO: 24

This gene is expressed primarily in fetal spleen.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, leukemia, lymphoma, AIDS, hematopoeitic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hematopoietic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., spleen and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of immune disorders including: leukemias, lymphomas, auto-immune, immuno-supressive (e.g. transplantation) and immunodeficiencies (e.g. AIDS) and hematopoietic disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 25

A closely related homolog of this gene was recently cloned by another group, calling the gene CDO, an oncogene-, serum-, and anchorage-regulated member of the Ig/fibronectin type III repeat family. (See Accession No. 2406628, and J. Cell Biol. 138(1): 203-213 (1997), herein incorporated by reference in its entirety.) Preferred polypeptide fragments comprise the amino acid sequence: FYIYYRPTDSDNDSDYKK DMVEGDKYWHSISHLQPETSYDIKMQCFNEGGESEFSNVMICETKARKSSGQP GRLPPPTLAPPQPPLPETIERPVGTGAMVARSSDLPYLIVGVVLGSIVLIIVTFIPF CLWRAWSKOKHTTDLGFPRSALPPSCPYTMVPLGGLPGHQAVDSPTSVASVD

10

15

20

25

GPVLM (SEQ ID NO:273); or YIYYRPTDSDNDSDYKKDMVEGDKYWHSISHLQ PETSYDIKMQCFNEGGESEFSNVMICETKARKS (SEQ ID NO:274).

This gene is expressed primarily in fetal lung and kidney, human embryo and osteoclastoma stromal cells and to a lesser extent in a variety of other tissues and cell types.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, developmental disorders and cancers, as well as pulmonary and renal disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the respiratory/pulmonary, skeletal and renal systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., lung, kidney, embryonic tissue, and bone cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder. relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:149 as residues: Thr-5 to Pro-18. Ala-76 to Thr-84.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of: osteoperosis, fracture, osteosarcoma, ossification, and osteonecrosis, as well as respiratory/pulmonary disorders, such as athesma, pulmonary edema, and renal conditions such as acute renal failure, kidney fibrosis and kidney tubule regeneration.

FEATURES OF PROTEIN ENCODED BY GENE NO: 26

This gene is homologous to the HIV envelope glycoprotein. (See Accession No. 2641463.) Preferred polypeptide fragments comprise the amino acid sequence: NVRALLHRMPEPPKINTAKFNNNKRKNLSL (SEQ ID NO:275).

This gene is expressed primarily in pineal gland and skin, and to a lesser extent in lung.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are

not limited to, neurological and behavior disorders; respiratory/pulmonary disorders, such as athesma, pulmonary edema; skin conditions such as eczema, psoriasis, acne and skin cancer, as well as AIDS. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous and respiratory systems, as well as skin and AIDS, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, pineal gland, epidermis, and pulmonary tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:150 as residues: Gln-15 to Gln-20.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of conditions which affect the above tissues, such as: skin cancer, eczema, psoriasis, acne, athesma, pulmonary edema, neuro-degenerative or developmental disorders such as Alzheimer's, depression, schizophrenia, dementia, and AIDS.

20

25

30

35

5

10

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 27

Preferred polypeptide encoded by this gene comprise the following amino acid sequence: NTNQREALQYAKNFQPFALNHQKDIQVLMGSLVYLRQGIENSPYVHIL LDANQWADICDIFTRDACALLGLSVESPLSVSFSAGCVALPALINIKAVIEQRQC TGVWNQKDELPIEVDLGKKCWYHSIFACPILRQQTTDNNPPMKLVCGHIISRD ALNKMFNGSKLKCPYCPMEQSPGDAKQIFF (SEQ ID NO:276). Polynucleotides encoding such polypeptides are also provided as are complementary polynucleotides thereto.

This gene is expressed primarily in liver (adult and fetal) and spleen tissue, and to a lesser extent in placenta, T helper cells, kidney tumor, ovarian tumor, melanocytes and fetal heart.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune and developmental diseases and disorders and liver diseases such as liver cancer. Similarly, polypeptides and antibodies directed to these

10

20

25

30

35

polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, circulatory and hematopoietic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., liver, spleen, placenta, blood cells, kidney, ovary and other reproductive tissue, melanocytes, and heart, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that the protein products of this gene are useful for study, diagnosis and treatment of growth, hematopoietic and immune system disorders particularly related to the liver.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 28

The translation product of this gene shares sequence homology with prostaglandin transporter which is thought to be important in metabolic and endocrine disorders. See, for example, Gastroenterology Oct:109(4):1274-1282 (1995). Preferred polypeptides encoded by this gene comprise the following amino acid sequence: SYLSACFAGCNSTNLTGCACLTTVPAENATVVPGKCPSPGCQEAFLTFLCVMCI CSLIGAMARHP (SEQ ID NO:277); and/or PSVIILIRTVSPELKSYALGVLFLLLRL LGFIPPPLIFGAGIDSTCLFWSTFCGEQGACVLYDNVVYRYLYVSIAIALKSFAFI (SEQ ID NO:278).

This gene is expressed primarily in hematopoietic and brain tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, metabolic, immune and endocrine diseases and disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the metabolic, immune and endocrine systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., endocrine tissue, hematopoietic tissue, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to

10

15

20

the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to prostaglandin (and anion) transporter indicates that polynucleotides and polypeptides corresponding to this gene are useful for study, diagnosis and treatment of endocrine, metabolic, immune and kidney disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 29

This gene is expressed primarily in early stage human lung.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, growth and respiratory disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the developmental and respiratory systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., pulmonary tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:153 as residues: Val-50 to Trp-55.

The tissue distribution indicates that the protein products of this gene are useful for study, diagnosis and treatment of respiratory and growth diseases and disorders.

25

30

35

FEATURES OF PROTEIN ENCODED BY GENE NO: 30

The translation product of this gene shares sequence homology with human DNA helicase which is thought to be important in accurate and complete DNA replication in creation of new cells. Preferred polypeptides encoded by this gene comprise the following amino acid sequence: QSLFTRFVRVGVPTVDLDAQGRARA SLCXXYNWRYKNLGNLPHVQLLPEFSTANAGLLYDFQLINVEDFQGVGESEPN PYFYQNLGEAEYVVALFMYMCLLGYPADKISILTTYNGQKHLIRDIINRRCGNN PLIGRPNKVTTVDRFQGQQNDYILLSLVRTRAVGHLRDVRRLVVAMSRAR (SEQ ID NO:279); and/or LVKEAKIIAMTCTHAALKRHDLVKLGFKYDNILMEE AAQILEIETFIPLLLQNPQDGFSRLKRWIMIGDHHQLPPVI (SEQ ID NO:280).

This gene is expressed primarily in testes tumor and to a lesser extent in adrenal

gland tumor and placenta.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, cancers and endocrine/growth disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine, developmental, and reproductive systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., testes and other reproductive tissue, adrenal gland, and placenta, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue-or-bodily fluid from an individual not having the disorder.

The tissue distribution and homology to DNA helicase indicates that the protein products of this gene are useful for study, treatment, and diagnosis of many cancer types, including testicular cancer; as well as disorders involving endocrine function and normal growth and development.

20

5

10

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 31

The translation product of this gene shares sequence homology with BID-apoptotic death gene (mouse), Genbank accession no. PID g1669514, which is thought to be important in programmed cell death.

25

30

35

This gene is expressed primarily in jurkat membrane bound polysomes and activated neutrophils and to a lesser extent in endothelial cells and human cerebellum.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, cancers and other proliferative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, endothelium, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma,

10

15

20

25

30

35

urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:155 as residues: Glu-4 to Leu-11, Cys-28 to Arg-35, Gln-50 to His-66, Glu-73 to Gln-79, Gly-94 to Ser-100, Arg-114 to Asp-126, Pro-139 to Lys-146.

The tissue distribution and homology to BID-apoptotic death gene indicates that the protein products of this gene are useful for study of cell death, and treatment and diagnosis of proliferative disorders and cancers. Apoptosis - programmed cell death - is a physiological mechanism involved in the deletion of peripheral T lymphocytes of the immune system, and its dysregulation can lead to a number of different pathogenic processes. Diseases associated with increased cell survival, or the inhibition of apoptosis, include cancers (such as follicular lymphomas, carcinomas with p53 mutations, and hormone-dependent tumors, such as breast cancer, prostrate cancer, Kaposiís sarcoma and ovarian cancer); autoimmune disorders (such as systemic lupus erythematosus and immune-related glomerulonephritis rheumatoid arthritis) and yiral infections (such as herpes viruses, pox viruses and adenoviruses), inflammation; graft vs. host disease, acute graft rejection, and chronic graft rejection. Diseases associated with increased apoptosis include AIDS; neurodegenerative disorders (such as Alzheimer's disease, Parkinson's disease, Amyotrophic lateral sclerosis, Retinitis pigmentosa, Cerebellar degeneration); myelodysplastic syndromes (such as aplastic anemia), ischemic injury (such as that caused by myocardial infarction, stroke and reperfusion injury), toxin-induced liver disease (such as that caused by alcohol), septic shock, cachexia and anorexia. Thus, the invention provides a method of enhancing apoptosis in an individual by treating the individual with a polypeptide encoded by this gene.

FEATURES OF PROTEIN ENCODED BY GENE NO: 32

The translation product of this gene shares sequence homology with human fructose transporter which is thought to be important in normal metabolic function and activity.

This gene is expressed primarily in T-cell lymphoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are

15

20

25

30

35

not limited to, leukemia and other cancers, and metabolic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoietic, lymph and metabolic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., T-cells and other blood cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:156 as residues: Pro-22 to Gly-48, Ser-54 to Pro-61.

The tissue distribution indicates that the protein products of this gene are useful for study of mechanisms leading to cancer, treatment and diagnosis of cancerous and pre-cancerous conditions; as well as the study and treatment of various metabolic diseases and disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 33

This gene is expressed primarily in human meningima.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, inflammation and other disorders of the CNS. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the CNS and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., meningima and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:157 as residues: Asn-23 to Pro-31.

The tissue distribution indicates that the protein products of this gene are useful for study, diagnosis and treatment of disorders of the CNS and inflammatory responses.

10

15

20

30

35

FEATURES OF PROTEIN ENCODED BY GENE NO: 34

This gene is expressed primarily in activated monocytes and wound healing tissues and to a lesser extent in fetal epithelium.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune and inflammatory disorders and wound healing and tissue repair dysfunctions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, epithelial and gastrointestinal systems, and healing wounds, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., monocytes and other blood cells, and epithelium, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:158 as residues: Ala-28 to Ala-33, Gly-35 to Glu-45.

The tissue distribution indicates that the protein products of this gene are useful for diagnosis, study and treatment of immune and inflammatory disorders and wound healing dysfunctions.

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 35

This gene is expressed primarily in human osteosarcoma and prostate cancer.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, skeletal and neoplastic conditions such as bone and prostate cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and skeletal systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., bone, and prostate, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial

10

15

20

25

30

35

fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:159 as residues: Ser-14 to Gly-22, Leu-37 to Gln-43.

The tissue distribution indicates that the protein products of this gene are useful for diagnosis and treatment of skeletal disorders and cancer.

FEATURES OF PROTEIN ENCODED BY GENE NO: 36

This gene encodes a protein which is highly homologous to a protein called congenital heart disease protein 5, presumably implicated in congenital heart disease (see Genbank PID g2810996).

This gene is expressed primarily in Hodgkin's lymphoma, erythroleukemia cells, and TNF activated synovial fibroblasts, to a lesser extent in ovarian cancer, cerebellum, spleen, fetal liver and placenta and finally to a lesser extent in various other mesenchymal tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, cancer, immune, hematopoietic and cardiovascular disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, hematopoietic and cardiovascular systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., heart and other cardiovascular tissue, lymphoid tissue, blood cells, bone marrow, ovary and other reproductive tissue, brain and other tissue of the nervous system, spleen, liver, and mesenchymal tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:160 as residues: Lys-41 to Met-49, Gln-54 to Glu-59, Glu-76 to Thr-88.

The homology of this gene and translation product to congenital heart disease protein 5 indicates a role for this protein in the diagnosis, prognosis and/or treatment of

WO 98/42738 PCT/US98/05311

34

heart disease or other cardiovascular related disorders. In addition, predominant expression in cells associated with the immune and hematopoetic system indicates a role for this protein in the treatment, diagnosis and/or prognosis of immune and autoimmune diseases, such as lupus, transplant rejection, allergic reactions, arthritis, asthma, immunodeficiency diseases, leukemia, AIDS, thymus disorders such as Graves 5 Disease, lymphocytic thyroiditis, hyperthyroidism and hypothyroidism, graft versus host reaction, graft versus host disease, transplant rejection, myelogenous leukemia, bone marrow fibrosis, and myeloproliferative disease. The protein could also be used to enhance or protect proliferation, differentiation and functional activation of hematopoietic progenitor cells such as bone marrow cells, which could be useful for 10 cancer patients undergoing chemotherapy or patients undergoing bone marrow transplantation. The protein may also be useful to increase the proliferation of peripheral blood leukocytes, which could be useful in the combat of a range of hematopoietic disorders including immunodeficiency diseases, leukemia, and septicemia.

15

20

25

FEATURES OF PROTEIN ENCODED BY GENE NO: 37

This gene is expressed primarily in ovarian cancer.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, urogenital neoplasias. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., ovary and other reproductive tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:161 as residues: Asn-22 to Asn-27.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for study, diagnosis and treatment of ovarian and other tumors.

30

10

15

20

25

30

35

FEATURES OF PROTEIN ENCODED BY GENE NO: 38

The translation product of this gene shares sequence homology with zinc finger proteins.

This gene is expressed primarily in various fetal, cancer, and endothelial lines.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune and growth disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cardiovascular system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., fetal tissue, and endothelial cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that the protein products of this gene are useful for study, diagnosis and treatment of immune and developmental conditions and cancer.

FEATURES OF PROTEIN ENCODED BY GENE NO: 39

This gene is expressed primarily in fetal, infant, and adult brain and to a lesser extent in other brain and endocrine organs and blastomas.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, brain tumors and neurodegenerative conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous and endocrine systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., brain and other tissue of the nervous system, endocrine tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the

WO 98/42738 PCT/US98/05311

36

disorder.

5

10

15

20

25

The tissue distribution indicates that the protein products of this gene are useful for the study, diagnosis and treatment of brain cancer and other neurological disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 40

The translation product of this gene shares sequence homology with vesicular glycoproteins and lectins. Preferred polypeptides encoded by this gene comprise the following amino acid sequence: DTYPNEEKQQERVFPXXSAMVNNGSLSYDHER DGRPTELGGCXAIVRNLHYDTFLVIRYVKRHLTIMMDIDGKHEWRDCIEVPGV RLPRGYYFGTSSITGDLSDNHDVISLKLFELTVERTPEEE (SEQ ID NO:281); and/or LKREHSLSKPYQGVGTGSSSLWNLMGNAMVMTQYIRLTPDMQSKQGA LWNRVPCFLRDWELQVHFKIHGQGKKNLHGDGLAIWYT (SEQ ID NO:282).

This gene is expressed primarily in infant brain and to a lesser extent in various normal and transformed neural, endocrine, and immune organs.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, neurological and neurodevelopmental conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous and hormonal systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., brain and other tissue of the nervous system, endocrine tissue, and tissue and cells of the immune system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:164 as residues: Pro-64 to Gly-71, Gly-94 to Leu-100, Thr-110 to Pro-116, Thr-135 to Arg-145, Glu-164 to Glu-171, Asp-204 to Asp-211, Arg-253 to His-261, Asn-312 to Tyr-323.

The tissue distribution indicates that the protein products of this gene are useful for the study, diagnosis and treatment of mental retardation and other neurological disorders and neoplasias.

30

10

15

20

FEATURES OF PROTEIN ENCODED BY GENE NO: 41

This gene displays homology to the glycosyltransferase family, which catalyze the addition of sialic acids to carbohydrate groups which are present on glycoproteins.

This gene is expressed primarily in smooth muscle and to a lesser extent in pineal gland, fetal liver, and infant brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, gastrointestinal injury, inflammatory and neurodegenerative conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and nervous systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues-(e.g., smooth muscle, pineal gland, liver, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:165 as residues: Ser-12 to Trp-21, Arg-24 to Pro-32, Asp-73 to Lys-82, Lys-90 to Ala-97.

The tissue distribution indicates that the protein products of this gene are useful for the study, diagnosis and treatment of neurodegenerative and growth disorders and gastrointestinal repair.

25

30

35

FEATURES OF PROTEIN ENCODED BY GENE NO: 42

The translation product of this gene shares sequence similarity with metallothionein polypeptides. See, for example, Proc. Natl. Acad. Sci. U S A 1992 Jul 15:89(14):6333-6337. Metallothioneins are believed to inhibit neuronal survival among other biological functions. Based on the sequence similarity (especially the conserved cysteine motifs characteristic of the metallothionein family) the translation product of this gene is expected to share certain biological activities with other members of the metallothionein polypeptide family. Preferred polypeptides encoded by this gene comprise the following amino acid sequence: PGTLQCSALHHDPGCANCSRFCRD CSPPACQC (SEQ ID NO:283).

This gene is expressed exclusively in placenta and fetal liver.

10

15

25

30

35

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, hematopoietic and immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., placenta, liver, brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to metallothionien indicates that the protein products of this gene are useful for diagnosis and treatment of immune and hematopoietic system disorders and neurological diseases, especially in fetal development.

20 FEATURES OF PROTEIN ENCODED BY GENE NO: 43

Preferred polypeptides encoded by this gene comprise the following amino acid sequence: FLYDVLMXHEAVMRTHQIQLPDPEFPS (SEQ ID NO:284).

This gene is expressed primarily in T-cells and synovial tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune system disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., synovial tissue, and T-cells and other blood cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

10

15

20

25

The tissue distribution indicates that the protein products of this gene are useful for treatment and diagnosis of disorders of the immune system.

FEATURES OF PROTEIN ENCODED BY GENE NO: 44

The translation product of this gene shares sequence similarity with several methyltransferases (e.g., see Genbank gil1065505).

This gene is expressed primarily in ovary, thymus, infant adrenal gland, tissues of the nervous system and the hematopoietic tissue, and to a lesser extent in adipose tissue and many other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, disorders of the reproductive system, the endocrine system, the hematopoietic system and the CNS. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, endocrine, CNS and reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., ovary and other reproductive tissue, thymus, adrenal gland, brain and other tissue of the nervous system, hematopoietic tissue, and adipose tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:168 as residues: Ser-3 to Gly-12, Asp-19 to Arg-31, Tyr-70 to Tyr-77, Asn-130 to Lys-140, Pro-165 to Gln-170, Pro-192 to Lys-199, Leu-216 to Glu-227, Glu-254 to Phe-281.

The tissue distribution and homology to methyltransferase indicates that the

protein products of this gene are useful for diagnosis and treatment of disorders of the

CNS, the hematopoietic system and reproductive organs and tissues. For example, the
abundant expression in the ovary may indicate that the gene product can be used as a
hormone with either systemic or reproductive functions; as growth factors for germ cell
maintenance and in vitro culture; as a fertility control agent; remedy for sexual

dysfunction or sex development disorders; diagnostics/treatment for ovarian tumors,
such as serous adenocarcinoma, dysgerminoma, embryonal carcinoma,

choriocarcinoma, teratoma, etc; The expression in thymus may indicate its utilities in T-cell development and thus its applications in immune related medical conditions, such as infection, allergy, immune deficiency, tissue/organ transplantation, etc.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 45

The translation product of this gene shares sequence homology with cytochrome C oxidase which is thought to be important in metabolic function of cells. This gene has now recently been published as estrogen response gene. See Genbank accession no. AB007618 and Mol. Cell. Biol. 18 (1), 442-449 (1998). See also J Immunol. Mar 1:154(5): 2384-2392 (1995), where the mouse homologue was published and implicated in siliocis.

This gene is expressed primarily in adipose tissue, kidney and fetal brain and to a lesser extent in several other tissues and organs.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, metabolic diseases involving especially adipose tissue, brain and kidney. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the CNS and vascular system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., adipose tissue, kidney, brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:169 as residues: Thr-5 to Ser-14.

The tissue distribution and homology to cytochrome C oxidase, estrogen response gene product and siliocis related gene product indicates that the protein products of this gene are useful for diagnosis and treatment of metabolic disorders in the CNS, adipose tissue and kidney, particularly siliocis.

FEATURES OF PROTEIN ENCODED BY GENE NO: 46

The translation product of this gene shares sequence homology with reticulocalbin. See, for example, J. Biochem. 117 (5), 1113-1119 (1995). Based on the

10

15

20

25

30

35

10

15

20

25

30

35

sequence similarity, the translation product of this gene is expected to share certain biological activities with reticulocalbin, e.g., Ca++ binding activities. This gene product is sometimes hereinafter referred to as "Reticulocalbin-2".

This gene is expressed primarily in breast, endothelial cells, synovial, heart and smooth muscle cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, diseases of the breast, vascular and skeletal/cardiac muscular system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the breast, vascular and skeleto-muscular system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., mammary tissue, endothelial cells, synovial tissue, heart and other cardiovascular tissue, and smooth muscle, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:170 as residues: Gly-16 to Arg-32, Ala-42 to Asn-50, Glu-66 to Gln-76, Arg-85 to Gly-94, Thr-108 to Asp-115, Trp-121 to Gly-130, Leu-137 to His-144, Glu-155 to Lys-161, Asp-175 to Ser-180, Glu-209 to Gly-217, Glu-232 to Glu-237, Thr-243 to Asp-261, Glu-287 to Arg-295.

The tissue distribution indicates that the protein products of this gene are useful for diagnosis and treatment of diseases of the vascular and skeletal/cardiac muscular system. The homology of the gene with reticulocalbin indicates its biological function in regulating calcium store, a particularly important function in muscular cell types. The gene expression in the heart may indicate its utilities in diagnosis and remedy in heart failure, ischemic heart diseases, cardiomyopathy, hypertension, arrhythmia, etc. The abundant expression in the breast may indicate its applications in breast neoplasia and breast cancers, such as fibroadenoma, papillary carcinoma, ductal carcinoma, Pagetís disease, medullary carcinoma, mucinous carcinoma, tubular carcinoma, secretory carcinoma and apocrine carcinoma; juvenile hypertrophy and gynecomastia, mastitis and abscess, duct ectasia, fat necrosis and fibrocystic diseases, etc.

10

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 47

The translation product of this gene shares weak sequence homology with H+transporting ATP synthase which is thought to be important in cell metabolism or signal transduction.

This gene is expressed only in testis.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of some types of diseases and conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain and hematopoietic tissues, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., testes and other reproductive tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Since only one out of about a million expressed sequence tag is found in testes indicates that its expression is low and selectively in testes. Since some of the genes only expressed in testes are usually expressed in brain or in certain induced hematopoietic cells/tissues, it is speculated that this gene to be expressed in brain or hematopoietic cells/tissues and is useful for diagnosis and treatment of disorders these systems.

25

30

35

DISCOUNT AND

20

FEATURES OF PROTEIN ENCODED BY GENE NO: 48

The translation product of this gene shares sequence homology with human polymeric immunoglobulin receptor (accession No.X73079) which is thought to be important in antibody recognition and immune defenses. In one embodiment, polypeptides of the invention comprise the sequence GWYWCG (SEQ ID NO:285). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in placenta and to a lesser extent in corpus callosum and fetal liver and spleen.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are

10

15

20

25

30

35

not limited to, disorders of the immune system, e.g. autoimmune diseases and immunodeficiency. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., placenta, liver, and spleen, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:172 as residues: Tyr-37 to Cys-49, Gly-51 to Tyr-56, Lys-88 to Trp-93, Leu-130 to Glu-136.

The tissue distribution and homology to human polymeric immunoglobulin receptor indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of immune disorders, e.g. autoimmune diseases and immunodeficiencies.

FEATURES OF PROTEIN ENCODED BY GENE NO: 49

This gene is expressed in thymus.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune disorder. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., thymus and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of immune disorders, e.g. autoimmunity and immunodeficiency.

10

15

20

25

FEATURES OF PROTEIN ENCODED BY GENE NO: 50

Preferred polypeptide encoded by this gene comprise the following amino acid sequence: MKVGARIRVKMSVNKAHPVVSTHWRWPAEWPQMFLHLAQEPRTE VKSRPLGLAGFIRQDSKTRKPLEQETIMSAADTALWPYGHGNREHQENELQKY LQYKDMHLLDSGQSLGHTHTLQGSHNLTALNI (SEQ ID NO:286). Polynucleotides encoding this polypeptide are also provided as are complementary polynucleotides thereto.

This gene is expressed primarily in adrenal gland, pituitary, T helper cells, and breast cells and to a lesser extent in a wide variety of tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the some diseases and conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and endocrine systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., adrenal gland, pituitary, T-cells and other blood cells, and mammary tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:174 as residues: Gln-39 to Ser-47, Arg-57 to Glu-67, Tyr-82 to Gln-95.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of a wide range of disorders, such as immune and endocrine disorders.

30 FEATURES OF PROTEIN ENCODED BY GENE NO: 51

The translation product of this gene shares sequence homology with human Sop2p-like protein which is important in cytoskeleton structure. In one embodiment, polypeptides of the invention comprise the sequence SLHKNSVSQISVLSGGKAKCS QFCTTGMDGGMSIWDVKSLESALKDLKI (SEQ ID NO:287). Polynucleotides encoding this polypeptide are also encompassed by the invention. This gene maps to chromosome 7. Therefore, polynucleotides of the invention can be used in linkage

35

10

15

analysis as a marker for chromosome 7.

This gene is expressed primarily in immune and hematopoietic tissues/cells and to a lesser extent in other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immunological and hematopoietic disorders and inflammation. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hematopoietic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., immune and hematopoietic tissue/cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:175 as residues: Lys-49 to Gln-54, Ala-61 to Arg-66, Lys-82 to Lys-87, Glu-126 to Val-133, His-136 to Ile-141, Glu-175 to Ser-187, Asp-286 to Leu-296, Ala-298 to Ser-310.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of immunological, hematopoietic, and inflammatory disorders, e.g, immunodeficiency, autoimmunity, inflammation.

25

30

35

20

FEATURES OF PROTEIN ENCODED BY GENE NO: 52

The translation product of this gene shares sequence homology with Caenorhabditis elegans R53.5 gene encoding a putative secreted protein without known function.

This gene is expressed primarily in endothelial cells, brain and several highly vascularized, and tumor tissues and to a lesser extent in other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, aberrant angiogensis and tumorigenesis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes

10

15

20

25

30

35

for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the vascular and brain system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., endothelial cells, brain and other tissue of the nervous system, and vascular tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:176 as residues: Thr-43 to Asn-60, Thr-106 to Phe-115, Asp-122 to Arg-133, Arg-186 to Asp-192, Leu-211 to Lys-216.

The tissue distribution and homology to a *C. elegans* secreted protein indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis or treatment of disorders in vascular or brain system, e.g. aberrant angiogenesis, ischemia, neurodegeneration, etc.

FEATURES OF PROTEIN ENCODED BY GENE NO: 53

In one embodiment, polypeptides of the invention comprise the sequence EASKSSHAGLDLFSVAACHRF (SEQ ID NO:288). Polynucleotides encoding this polypeptide are also encompassed by the invention.

This gene is expressed primarily in T-cells and to a lesser extent in brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, lymphocytic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the lymphoid system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., T-cells and other blood cells, brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:177 as residues: Pro-3 to Thr-8, Arg-37 to Asp-46.

15

20

25

30

35

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis, treatment, and cure of lymphocytic disorders.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 54

The translation product of this gene shares sequence homology with secreted cartilage matrix protein, a major component of the extracellular matrix of nonarticular cartilage which is thought to be important in cartilage structure. In specific embodiments, polypeptides of the invention comprise the sequence: RCKKCTEGPI DLVFVIDGSKSLGEENFEVVKQF (SEQ ID NO:297); VTGIIDSLTISPKAARVGL LQYSTQVH (SEQ ID NO:290); TEFTLRNFNSAKDMKKAVAHMKYM (SEQ ID NO:291); GKGSMTGLALKHMFERSFTQGEGARPF (SEQ ID NO:292); STRVP RAAIVFTDGRAQDDVSEWASKAKANGITMYAVGVGKAIE (SEQ ID NO:293); EELQEIASEPTNKHLFYAEDFSTMDEISEKLKKGICEALEDS (SEQ ID NO:294); TQRLEEMTQRM (SEQ ID NO:295); PQGCPEQPLH (SEQ ID NO:296); and/or YMGKGSMTGLALKHMFERSFT (SEQ ID NO:289). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in placenta, infant brain, prostate, fetal lung and to a lesser extent in endometrium and fetal tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, abnormal placenta and pregnancy, disorder and injury in brain, prostate, and vasculature. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproduction, neuronal, and vascular systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., placenta, brain and other tissue of the nervous system, prostate, lung and endometrium, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to cartilage matrix protein indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis,

10

15

20

25

30

35

treatment, and cure of abnormalities in placenta and pregnancy, disorder and injury in brain, prostate, and vasculature.

FEATURES OF PROTEIN ENCODED BY GENE NO: 55

The translation product of this gene is the human ortholog of bovine and hamster CII-3, a succinate-ubiquinone oxidoreductase complex II membrane-intrinsic subunit, which is thought to be important in mitochondrial electron transport chain during metabolism. In specific embodiments, the polypeptides of the invention compriseMAALLLRHVGRHCLRAHFSPQLCIRNAVPLGTTAKEEMERFWNKNIG SNRPLSPHITIYS (SEQ ID NO:298); VFPLMYHTWNGIRHLMWDLGKGLKIPQL YQSG (SEQ ID NO:299); MAALLLRHVGRHCLRAH (SEQ ID NO:300); VKSLCL GPALIHTAKFAL (SEQ ID NO:301); VFPLMYHTWNGIRHLMWDLGKGL (SEQ ID NO:302).

This gene is expressed in 8-week old early stage human.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, metabolism disorder. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the [insert system where a related disease state is likely, e.g., immune], expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis, treatment, and cure of metabolism disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 56

This gene is expressed primarily in umbilical vein endothelial cells, human ovarian tumor cells, human meningima cells, and human Jurkat membrane bound polysomes. In specific embodiments, polypeptides of the invention comprise the amino acid sequence: RVWDVRPFAPKERCVKIFQGNV (SEQ ID NO:303); HNFEKNLL

10

15

20

RCSWSPDGSKIAAGSADRFVYV (SEQ ID NO:304); and/or WDTTSRRILYKLPG HAGSINEVAFHPDEPI (SEQ ID NO:305). Polynucleotides encoding these polypeptides are also encompassed by the invention.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, inflammation, immune and cardiovascular disorders and urogenital neoplasias. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of these tissue(s) or cell type(s). For a number of disorders of the above tissues or cells. particularly of the immune, neurological, urogenital, reproductive system and vascular systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, cells, endothelial cells, ovary -and other-reproductive-tissue, meningima, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:143 as residues: Phe-71 to Arg-76, Pro-82 to His-87, Glu-103 to Ala-111.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of immune disorders including: leukemias, lymphomas, auto-immune, immuno-supressive (e.g. transplantation) and immunodeficiencies (e.g. AIDS) and hematopoietic disorders. In addition, expression in ovarian tumor cells suggests that polynucleotides and polypeptides corresponding to this gene are useful for study, diagnosis, and treatment of ovarian tumors, and other tumors and neoplasias. Further, endothelial cell expression suggests a role in cadiovascular or respiratory/pulmonary disorders or infections (athsma, pulmonary edema, pneumonia).

30

35

25

FEATURES OF PROTEIN ENCODED BY GENE NO: 57

The translation product of this gene shares sequence homology with type I collagen. In specific embodiments, the polypeptides of the invention comprise the sequence: GRIPAPAPSVPAGPDSR (SEQ ID NO:309); VRGRTVLRPGLDAEPE LSPE (SEQ ID NO:306); EQRVLERKLKKERKKEERQ (SEQ ID NO:307); ARRSG

10

15

AELAWDYLCRWAQKHKNWRFQKTRQTWLLLHMYDSDKVPDEHFSTLLAYLE GLQGR (SEQ ID NO:255); and/or RLREAGLVAQHPP (SEQ ID NO:308).

Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in epididymus, prostate cell line (LNCAP), and pituitary gland; and to a lesser extent in many other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, abnormalities of the epididymus, prostate (especially prostate cancer), and pituitary gland. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the male reproductive system and neuroendocrine system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., epididymus and other reproductive tissue, prostate, and pituitary gland, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to type I collagen, indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of abnormalities of the epididymus, prostate (especially prostate cancer), and pituitary gland.

25

30

35

20

FEATURES OF PROTEIN ENCODED BY GENE NO: 58

This gene is expressed primarily in the frontal cortex of the brain from a schizophrenic individual.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, schizophrenia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., brain and other tissue of

20

the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of schizophrenia.

FEATURES OF PROTEIN ENCODED BY GENE NO: 59

The polypeptide encoded by Gene 59 is homologous to human surface 4 integral membrane protein. In specific embodiments, the polypeptides of the invention comprise the sequence: TGCVLVLSRNFVQYACFGLFGIIALQTIAYSILWDLKF LMRN (SEQ ID NO:310); SRSEGKSMFAGVPTMRESSPKQYMQLGGRVLLV LMFMTLLHFDASFFSIVQNIVG (SEQ IDNO:311); GTAEDFADQFLRVTKQYLP HVARLCLISTFLEDGIRMFQWSEQRDYIDTTWNCGYLLAS (SEQ ID NO:312); LMRNESRS (SEQ ID NO:314); ASFLLSRTSWGTA (SEQ ID NO:315); and/or ASFLLSRTSWGTALMIL (SEQ ID NO:313). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in Hodgkin's lymphoma and lung; and to a lesser extent in many other human tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, Hodgkin's lymphoma, tumors or other abnormalities of the lung. 25 Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and respiratory systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., lymphoid tissue, and pulmonary tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, 30 plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID 35 NO:183 as residues: Met-20 to Trp-27.

The tissue distribution indicates that polynucleotides and polypeptides

10

15

20

25

30

35

corresponding to this gene are useful for diagnosis and treatment of Hodgkin's lymphoma, tumors or other abnormalities of the lung.

FEATURES OF PROTEIN ENCODED BY GENE NO: 60

This gene is expressed primarily in bone cancer and stomach cancer, and to a lesser extent in many other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, bone cancer and stomach cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the bone, and the stomach, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., bone, and stomach, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of bone cancer and stomach cancer and possibly other cancers.

FEATURES OF PROTEIN ENCODED BY GENE NO: 61

This gene is expressed primarily in epididymus, and lymph node of breast cancer, and to a lesser extent in many other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, abnormalities of the epididymus, and breast cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the epididymus and breast, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., epididymus and other reproductive tissue, lymphoid tissue, and mammary tissue, and cancerous and wounded tissues) or bodily fluids

15

20

25

30

35

(e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:185 as residues: Arg-57 to Ser-65.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of abnormalities of the epididymus, and breast cancer.

10 FEATURES OF PROTEIN ENCODED BY GENE NO: 62

The translation product of this gene appears to be the human homolog of bovine NADH dehydrogenase which is thought to be important in cellular metabolism. In specific embodiments, the polypeptides of the invention comprise the amino acid sequence: SMSALTRLASFARVGGRLFRSGCARTAGDGGVRHAGGGVHIEPRY RQFPQLTRSQVFQSEFFSGLMWFWILWRFWHDSEEVLGHFPYPDPSQWTDEEL GIPPDDED (SEQ ID NO:323), or fragments thereof. Polynucleotides encoding this polypeptide are also encompassed by the invention.

This gene is expressed in larynx tumor, lymph node, brain amygdala, human cardiomyopathy, and retina.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, diseases affecting cellular metabolism. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., larynx, lymphoid tissue, brain and other tissue of the nervous system, heart and cardiovascular tissue, and retina, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:208 as residues: Pro-27 to Gln-32, Arg-42 to Glu-51.

The tissue distribution and homology to NADH dehydrogenase indicates that

10

15

20

polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of diseases involving cellular metabolism.

FEATURES OF PROTEIN ENCODED BY GENE NO: 63

This gene is expressed primarily in amygdala, and to a lesser extent in many other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, abnormalities of the amygdala. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the amygdala, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., amygdala, and lymphoid tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:187 as residues: Gln-17 to Glu-29, Pro-41 to Phe-46, Ser-59 to Ile-70, Thr-97 to Leu-105.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of abnormalities of amygdala.

25

30

35

FEATURES OF PROTEIN ENCODED BY GENE NO: 64

This gene is expressed primarily in female bladder, and to a lesser extent in chronic synovitis and hemangiopericytoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, bladder cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the urinary tract, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., bladder, synovial tissue, and

vascular tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:188 as residues: Pro-2 to Gln-7, Pro-27 to Phe-34.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatments of defects of the urinary tract, especially bladder cancer.

10

15

20

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 65

This gene is expressed primarily in fetal spleen, and to a lesser extent in hemangiopericytoma, thymus, and synovial sarcoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, defects of immune of hematopoietic systems. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune of hematopoietic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., spleen, vascular tissue, thymus, blood cells, and synovial tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The protein product of this gene is useful for treatment of defects of the immune or hematopoietic systems, because of the gene's expression in thymus and spleen.

30

35

25

FEATURES OF PROTEIN ENCODED BY GENE NO: 66

This gene is expressed primarily in human pituitary and to a lesser extent in placenta and fetal lung.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are

10

15

20

25

30

35

not limited to, endocrine growth disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., pituitary and other endocrine tissue, placenta, and pulmonary tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:190 as residues: Val-38 to Asn-44, Gly-53 to Ser-65.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment of growth disorders related to pituitary dysfunction.

FEATURES OF PROTEIN ENCODED BY GENE NO: 67

The translation product of this gene shares sequence homology with a Caenorhabditis elegans gene of unknown function. In specific embodiments, the polypeptides of the invention comprise the sequence: DPRRPNKVLRYKPPPSE CNPALDDPTP (SEQ ID NO:317); DYMNLLGMIFSMCGLMLKLKWCAWVA VYCS (SEQ ID NO:318); FISFANSRSSEDTKQMMSSF (SEQ ID NO:316); and/or MLSISAVVMSYLQNPQPMTPPW (SEQ ID NO:319). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in primary breast cancer and lymph node breast cancer and to a lesser extent in adult brain, lung cancer, colon cancer, epithelioid sarcoma, and Caco-2 cell line.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cancer and tumor tissues, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., mammary tissue, lymphoid tissue, brain and other tissue of the nervous system, lung, colon, and

10

15

20

25

30

35

epithelium, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:191 as residues: Asn-34 to Lys-42.

The tissue distribution in a variety of cancer tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of a variety of cancer and tumor types.

FEATURES OF PROTEIN ENCODED BY GENE NO: 68

The translation product of this gene shares sequence homology with steroid membrane binding protein. The translation product of this gene has recently been published as progesterone binding protein. See Genbank AJ002030. Preferred polypeptides encoded by this gene comprise the following amino acid sequence: AAGDGDVKLGTLGSGSESSNDGGSESPGDAGAAAXGGGWAAAALALLTG GGE (SEQ ID NO:320).

This gene is expressed primarily in breast, and to a lesser extent in placenta and fetal tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, breast cancer or developmental disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of breast or fetal tissues, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., mammary tissue, placenta, and fetal tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:192 as residues: Pro-43 to Asp-49, Gln-54 to Pro-64, Asp-110 to Asp-118, Lys-138 to Tyr-143, Pro-150 to Asp-170.

The tissue distribution and homology to steroid membrane binding protein and to progesterone binding protein indicates that the protein products of this gene are

useful for treatment of breast cancers, especially those caused by estrogen and progesterone binding.

FEATURES OF PROTEIN ENCODED BY GENE NO: 69

5

10

15

20

25

30

35

Preferred polypeptides encoded by this gene comprise the following amino acid sequence: AADNYGIPRACRNSARSYGAAWLLLXPAGSSRVEPTQDISISDQLGG QDVPVFRNLSLLVVGVGAVFSLLFHLGTRERRRPHAXEPGEHTPLLAPATAQPL LLWKHWLREXAFYQVGILYMTTRLIVNLSQTYMAMYLTYSLHLPKKFIATIPLV MYLSGFLSSFLMKPINKCIGRN (SEQ ID NO:321).

This gene is expressed primarily in macrophage (GM-CSF treated), and to a lesser extent in monocytes and dendritic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, inflammation and infection. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., macrophages and other blood cells, and dendritic cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that the protein products of this gene are useful for treatment of infection or inflammation or other events or defects involving the immune system.

FEATURES OF PROTEIN ENCODED BY GENE NO: 70

This gene is expressed primarily in adult brain and to a lesser extent in thyroid, 12 week old early stage human, and stromal cell TF274.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, neurological or neuro-endocrine diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes

10

20

25

30

35

for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous or endocrine systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system, thyroid, and stromal cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:194 as residues: Pro-65 to Cys-71.

The tissue distribution indicates that the protein products of this gene are useful for treatment and diagnosis of neurological diseases or metabolic conditions involving the neuro-endocrine system.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 71

This gene is expressed in T-cell helper and to a lesser extent in adult brain and adult testes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune disorders, meningitis or reproductive problems. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, neural and reproductive systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., T-cells and other blood cells, brain and other tissue of the nervous system, testes and other reproductive tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:195 as residues: Val-18 to Tyr-24, Ala-89 to Asp-99, Asp-104 to Ala-117, Leu-121 to Pro-136.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis immune and

10

15

20

25

30

reproductive disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 72

The translated polypeptide of this contig has a high degree of identity with the Ob Receptor-Associated Protein deposited as GenBank Accession No. 2266638. No function has been determined for the Ob Receptor-Associated Protein, however it is expressed upon stimulation of the Ob Receptor by Leptin.

This gene is expressed in T-cells and to a lesser extent in endothelial and bone marrow cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, acute lymphoblastic leukemia, hematapoetic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hematapoetic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., T-cells and other blood cells, endothelial cells, and bone marrow, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 196 as residues: Ser-61 to Trp-70.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of leukemia and other disorders of the primary immune system. In addition, since this gene appears to be related to the Ob Receptor-Related Protein, it is likely that this polypeptide is also involved in the Ob/Leptin signal transduction cascade. As a result, this protein may be of use in the molecular diagnosis and therapeutic intervention of obesity and related disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 73

The translation product of this contig has homology with furin, a protein thought to be a key endopeptidase in the constitutive secretory pathway. The identification and initial characterization of Furin was reported by Takahasi and

10

15

20

colleagues (Biochem Biophys Res Commun 1993 Sep 15;195(2):1019-1026).

This gene is expressed in neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, diseases of the immune system such as allergies, wound healing and antigen recognition. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., neutrophils and other blood cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment of allergies or other immune disorders since neutrophils are an important part of an allergic response. Further, since this protein appears to be related to Furin, it can be used diagnostically and therapeutically to treat secretory protein processing disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 74

This gene is expressed in the frontal cortex.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, of the motor activity and sensory functions that involve the central nervous system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene

expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of neural disorders that affect cognitive functions.

FEATURES OF PROTEIN ENCODED BY GENE NO: 75

The translation product of this gene shares sequence homology with inorganic pyrophophatase which is thought to be important in the catalysis the hydrolysis of diphosphate bonds, chiefly in nucleoside di- and triphosphates and essential enzymes that are important for controlling the cellular levels of inorganic pyrophosphate (PPi). The bovine homolog of this gene has been identified by Yang and Wensel (J. Biol. Chem. 267:24641-24647 (1992)).

This gene is expressed in osteoclastoma cells and to a lesser extent in epithelial cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, osteoporosis and other bone weakening diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skeletal system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., bone, and epithelial cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:199 as residues: Lys-22 to Tyr-28, Asp-64 to Lys-77, Pro-86 to Ile-91, Gln-99 to Pro-119, Tyr-169 to Asp-174, Lys-176 to Gly-181, Trp-189 to Asn-202, Lys-233 to Gly-239, Ser-250 to Asp-257.

The tissue distribution and homology to inorganic pyrophophatase indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of osteoporosis through the removal of bone by demineralization.

30

5

10

15

20

25

10

15

20

25

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 76

The translation product of this gene shares exact sequence homology with ATP sulfurylase/APS kinase (GenBank Accession No. 2673862) which is thought to be important in biosynthesis of the activated sulfate donor, adenosine 3'-phosphate 5'-phosphosulfate, involves the sequential action of two enzyme activities: ATP sulfurylase, which catalyzes the formation of adenosine 5'-phosphosulfate (APS) from ATP and free sulfate, and APS kinase, which subsequently phosphorylates APS to produce adenosine 3'-phosphate 5'-phosphosulfate.

This gene is expressed in osteoclastoma cells and to a lesser extent in developmental tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, antibiotic resistant bacterial infections, osteoarthritis and other auto immune diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune or skeletal structure expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., bone, and developmental tissues, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:200 as residues: Asn-15 to Trp-20, Ser-36 to Gly-41, Pro-103 to Val-110, Pro-134 to Arg-143, Leu-173 to Arg-178, Ser-190 to Ala-197, His-314 to Arg-319, Arg-354 to Asn-362, Asp-391 to Arg-397, Glu-402 to Asp-409, Asp-434 to Leu-439, Glu-441 to Arg-446, Gly-455 to Asp-462, Pro-528 to His-541, Asn-566 to Arg-571, Tyr-574 to Glu-581, Thr-589 to Glu-603.

The tissue distribution and homology to ATP sulfurylase/APS kinase indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment or detection of autoimmune diseases.

FEATURES OF PROTEIN ENCODED BY GENE NO: 77

This polypeptide is identical to the SLP-76-associated protein reported by Musci and colleagues (J. Biol. Chem. 272 (18), 11674-11677 (1997)) and to the FYB protein

10

15

20

25

30

35

reported by da Silva and coworkers (Proc. Natl. Acad. Sci. U.S.A. (1997) In press). These proteins have been reported to be novel T-cell Proteins which bind FYN and SLP-76 and regulate IL-2 production. Preferred polypeptides encoded by this gene comprise the following amino acid sequence: RITDNPEGKWLGRTARGSYGYIK TTAVEIXYDSLKLKKDSLGAPSRPIEDDQEVYDDVAEQDDISSHSQSGSGGIFPP PPDDDIYDGIEEEDADDGFPAPPKQLDMGDEVYDDVDTSDFPVSSAEMSQGTNV GKAKTEEKDLKKLKKQXKEXKDFRKKFKYDGEIRVLYSTKVTTSITSKKWGT RDLQVKPGESLEVIQTTDDTKVLCRNEEGKYGYVLRSYLADNDGEIYDDIADGC IYDND (SEQ ID NO:322).

This gene is expressed in CD34 positive cells (hematopoietic progenitor cells) and to a lesser extent in adult spleen derived from a chronic lymphocytic leukemia patient.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, chronic lymphocytic leukemia; hematopoietic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hematopoietic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., T-cells and other blood cells, bone marrow, hematopoietic cells, and spleen, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Further, nucleic acids and polypeptides of the present invention are useful both diagnostically and therapeutically in the intervention of immune and other disorders in which the ability to alter IL-2 expression is desired. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:201 as residues: Ala-17 to Lys-37, Val-39 to Ser-45, Lys-59 to His-70, Arg-90 to Leu-95, Lys-97 to Lys-107, Ser-117 to Leu-124, Phe-133 to Ser-138, Trp-146 to Leu-167, Pro-175 to Asn-185, Lys-190 to Ser-211, Pro-213 to Ser-222, His-230 to Pro-235, Pro-240 to Pro-246, Pro-253 to Gly-261, Leu-271 to Leu-303, Leu-305 to Leu-326, Lys-343 to Leu-349, Thr-363 to Leu-371, Arg-373 to Tyr-381, Tyr-391 to Leu-401, Pro-404 to Val-414, Ser-426 to Ser-432, Ile-448 to Ser-457, Gln-462 to Trp-468, Lys-477 to Ser-501, Asp-518 to Ser-523, Ala-541 to Gln-554.

15

20

25

30

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment of a variety of hematopoietic disorders. The noted expression of this gene in the hematopoietic progenitor cell compartment - as determined by its expression on CD34 positive hematopoietic stem and progenitor cells - indicates that it plays a critical role in the expansion or proliferation of hematopoietic stem/progenitor cells, as well as in the differentiation of the various blood cell lineages. Thus it could be useful in the reconstitution of the hematopoietic system of patients with leukemias and other hematopoietic diseases.

10 FEATURES OF PROTEIN ENCODED BY GENE NO: 78

This gene is homologous to heparin cofactor II (HCII) which is a 66-kDa plasma glycoprotein that inhibits thrombin rapidly in the presence of dermatan sulfate or heparin.

This gene is expressed in apoptotic and anergic T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, thrombopienia T-cell lymphomas; Hodgkin's lymphoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system most notably the T-cell compartment, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., T-cells and other blood cells, and lymphoid tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The homology to heparin cofactor II (HCII) and the tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic disorders particularly in thrombopoesis, most notably of the T-cell compartment. This could include immune modulation, inflammation, immune surveillance, graft rejection, and autoimmunity.

35 FEATURES OF PROTEIN ENCODED BY GENE NO: 79

The translation product of this gene shares sequence homology with a mouse

10

15

20

protein believed to represent an integral membrane protein.

This gene is expressed in fetal cochlea and epididymus and to a lesser extent in adult spleen and osteoclastoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, osteoclastoma; disorders of the inner ear; male fertility disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the inner ear; male reproductive tract; bone; and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., cochlea, epididymus and other reproductive tissue, spleen, and bone, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:203 as residues: Lys-13 to Gly-23, Cys-38 to Asp-43, Gly-48 to Trp-53, Cys-223 to Ile-237, Ile-240 to Ser-246.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment of hearing and fertility disorders. Likewise, it may have a role in the modulation of immune function and in the treatment of osteoporosis.

25

35

FEATURES OF PROTEIN ENCODED BY GENE NO: 80

The translation product of this gene shares sequence homology with reticulocalbin which is thought to be important in the binding of calcium, particularly within the endoplasmic reticulum.

This gene is expressed in endothelial cells and stromal cells and to a lesser extent in osteoblasts, osteoclasts, and T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, osteoperosis; osteoclastomas; T-cell lymphomas; Hodgkin's disease. Similarly, polypeptides and antibodies directed to these polypeptides are useful in

10

15

20

25

30

35

providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the vasculature, bone, and immune systems - particularly the T-cell compartments, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., endothelial cells, stromal cells, bone, T-cells and other blood cells, and lymphoid tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:204 as residues: Lys-20 to Arg-27, Pro-32 to Asp-48, Leu-64 to Arg-72, Asp-108 to Lys-114, Glu-128 to Thr-133, Asp-139 to Phe-147, Thr-196 to Ala-204, Tyr-218 to Glu-228, Val-230 to Gln-236, Arg-241 to Lys-255, Glu-276 to Lys-287.

The tissue distribution and homology to reticulocalbin indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of bone disorders such as osteoporosis; the diagnosis and treatment of T-cell lymphomas and Hodgkin's lymphoma; and the treatment of diseases and defects of the vasculature, such as vascular leak syndrome and aberrant angiogenesis that accompanies tumor growth.

FEATURES OF PROTEIN ENCODED BY GENE NO: 81

The translation product of this gene shares sequence homology with a family of peptide transport genes - particularly the AtPTR2-B gene from *Arabidopsis* - which are thought to be important in the uptake of small peptides.

This gene is expressed in a number of fetal tissues, most notably lung, brain, cochlea, and liver/spleen, and to a lesser extent in osteoclastoma and endometrial tumors.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, osteoclastoma; endometrial tumors; cancer; leukemias. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the bone and endometrium, expression of this gene at significantly higher or lower levels may be

10

15

20

25

30

35

routinely detected in certain tissues (e.g., fetal tissue, pulmonary tissue, bone, brain and other tissue of the nervous system, cochlea, liver, and spleen, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:205 as residues: Lys-186 to Asn-199, Pro-202 to Ala-207.

The tissue distribution and homology to peptide transport genes indicates that polynucleotides and polypeptides corresponding to this gene are useful for the control of cell proliferation, owing to its strong expression in fetal tissues undergoing active cell division, as well as its expression in a variety of tumors or cancers of adult tissues. Potentially, it may regulate the uptake of peptides that stimulate cell proliferation. This gene product may also be useful in stimulating the uptake of a variety of peptide-based drug compounds.

FEATURES OF PROTEIN ENCODED BY GENE NO: 82

This gene is expressed in fetal liver and spleen and to a lesser extent in endothelial cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, cancer and tumors of a hematopoietic and/or endothelial cell origin; leukemias. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and/or vasculature, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., liver, spleen, endothelial cells, vascular tissue, and tissue and cells of the immune system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:206 as residues: Met-1 to Asp-9, Arg-66 to Gly-76, Asp-164 to Arg-171.

The tissue distribution indicates that polynucleotides and polypeptides

10

15

20

25

30

corresponding to this gene are useful for the treatment of disorders of the immune system. Expression of this gene product in both fetal liver/spleen and endothelial cells indicates that it may be expressed in the hemangioblast, the progenitor cell for both the immune system and the vasculature. Thus, it is most likely expressed in hematopoietic stem cells, and may be useful for the expansion of hematopoietic stem and progenitor cells in conjunction with cancer treatment for a variety of leukemias.

FEATURES OF PROTEIN ENCODED BY GENE NO: 84

The translation product of this gene shares sequence homology with NADH dehydrogenase which is thought to be important in cellular metabolism.

This gene is expressed in fetal dura mater and to a lesser extent in T-cells and hypothalamus.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, diseases affecting cellular metabolism. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., fetal tissue, T-cells and other blood cells, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:208 as residues: Pro-27 to Gln-32, Arg-42 to Glu-51.

The tissue distribution and homology to NADH dehydrogenase indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of diseases involving cellular metabolism.

10

15

20

25

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 85

The translation product of this gene shares sequence homology with I-TRAF, a novel TNF receptor associated factor (TRAF)-interacting protein that regulates TNF receptor-mediated signal transduction. This protein is thought to be important in regulating the cellular response to tumor necrosis factor (TNF), which is an important mediator of inflammation.

This gene is expressed in endothelial cells and to a lesser extent in glioblastoma and osteoblastoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, inflammation; glioblastoma and osteoblastoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., endothelial cells, bone, and glial cells and tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:209 as residues: Glu-15 to Thr-22, Glu-46 to Leu-62, Arg-103 to Glu-119, Gln-127 to Glu-132, Asn-152 to Trp-158, Gln-191 to Gln-210, Glu-264 to Thr-271, Tyr-282 to Leu-288, Trp-319 to Thr-331, Glu-335 to Ser-348, Ser-353 to Ser-358, Asp-382 to Asn-392.

The tissue distribution and homology to I-TRAF indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of inflammatory diseases, including rheumatoid arthritis, sepsis, inflammatory bowel disease, and psoriasis, particularly where tumor necrosis factor is known to be involved.

10

15

20

30

35

FEATURES OF PROTEIN ENCODED BY GENE NO: 86

This gene has homology with a candidate gene involved in X-linked. Retinopathy reported by Wong and colleagues (Genomics 15:467-471 (1993)).

This gene is expressed in a T-cell line.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, inflammation and autoimmune diseases; T-cell lymphoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., T-cells and other blood cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of inflammatory disorders such as sepsis, inflammatory bowel disease, psoriasis, and rheumatoid arthritis as well as autoimmune disease such as lupus. It could also be useful in immune modulation and in the process of immune surveillance. The present invention can be used diagnostically and therapeutically to treat X-linked Retinopathy.

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 87

This gene is expressed in human brain tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, brain disorders; neurodegenerative disorders; tumors of a brain origin. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma,

10

urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:211 as residues: Cys-32 to Tyr-38.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of CNS disorders such as epilepsy, paranoia, depression, Alzheimer's disease, and schizophrenia. It could be useful in the survival and/or proliferation of neurons and could effect neuronal regeneration.

Last AA of ORF	30	4	69	18,	38	22	109
First AA of Secreted Portion		. 27	45	. 56	25	23	21
First Last AA AA of of of Sig Sig Pep Pep		26	<u>4</u> '	25		22	20
First AA of Sig Pep		-		- I	-	-	1
\$8 BBB>	125	126	212	213	127	128	129
of AA I First SEQ AA of ID Signal NO:	353	128	170	413	8	006	103
5° NT of Start Codor	353	128	170	413	66	006	103
3' NT of Clone Seq.	1607	1786	1487	1637	1212	2061	733
5' NT 3' NT of Of Clone Clone Seq.	247		79	394	-	882	10
Total NT Seq.	1679	1830	1487	1653	1212	2061	1412
X Seo	Ţ	12	86	- - -	13	14	15
Vector	Uni-ZAP XR	Uni-ZAP XR	pBluescript	Uni-ZAP XR	Uni-ZAP XR	Uni-ZAP XR	Uni-ZAP XR
ATCC Deposit Nr and Date	97923 03/07/97 209071 05/22/97	97923 03/07/97 209071 05/22/97	xxxxx 03/19/98	209641 02/25/98	97923 03/07/97 209071 05/22/97	97923 03/07/97 209071 05/22/97	97923 03/07/97 209071 05/22/97
cDNA Clone ID	HAGEW82					HAPBL78	HASAV70
Gene No.	I	2	2	. 2	3	4	\$

st F KF	62	. 529	2	9	2	∞
Last AA of ORF	9	2	52	99	215	48
First AA of Secreted Portion	18		24	- 18	19	27
First Last AA AA Iof of Sig Sig Pep	17		23	17	18	26
First AA of Sig Pep	1		1	ľ	Ī	1
¥ŠĐŠ¥ ׊Њ	130	131	132	-	134	135
of AA F First SEQ AA of D Signal NO: Pep Y	538	181	98	192	401	793
of of Start Codon	538	181	98	192	401	793
S' NT 3' NT of Ol Clone Clone Seq. Seq.	088		1007	1393	1070	2011
S' NT 3' NT of of Clone Clone Seq. Seq.	276	1	98	132	772	614
Total NT Seq.	1052	683	1054	1393	1215	2042
× Še Še X	16	17	<u>8</u>	19	20	21
Vector	Uni-ZAP XR	Uni-ZAP XR	Uni-ZAP XR		Uni-ZAP XR	Uni-ZAP XR
ATCC Deposit Nr and Date	97923 03/07/97 209071 05/22/97	97923 03/07/97 209071 05/22/97	97923 03/07/97 209071 05/22/97	97923 03/07/97 209071 05/22/97	97923 03/07/97 209071 05/22/97	97923 03/07/97 209071 05/22/97
cDNA Clone ID	HBNAF22	HBNBL77	HCDDR90	HCEEF50	HCEMU42	HCENE16
Gene No.	9	7	∞	6	10	11

# J. L		T	To .				
Last AA of	29	51	539	8	26	48	200
First AA of Secreted Portion	. 24	30	- 31	23	27	37	78
First Last AA AA of of Sig Sig Pep Pep	23	29	30	22	26	36	27
		-	1	-	-	-	-
¥SÖÐS;¥	136	137	138	214	139	215	140
S' NT of AA First SEQ AA of ID Signal NO: Pep Y	69	68	808	515	196	295	02
5' NT of Start Codon	69	68	808	515	196	295	70
S' NT 3' NT of of of Clone Clone Seq. Seq.	1872	289	3532	1115	2007	734	717
5' NT of Clone Seq.	21	1	2821	435	171	25	-
Total NT Seq.	1872	289	3533	1145	1148	734	717
NA NO:		_23	24	<u>1</u>	25	101	26
Vector	Uni-ZAP XR		Uni-ZAP XR	Uni-ZAP XR	pBluescript	Uni-ZAP XR	Uni-ZAP XR
ATCC Deposit Nr and Date	97923 03/07/97 209071 05/22/97	97923 03/07/97 209071 05/22/97	97923 03/07/97 209071 05/22/97	97923 03/07/97 209071 05/22/97	209179 07/24/97	97923 03/07/97 209071 05/22/97	97923 03/07/97 209071 05/22/97
	HMSJJ74		_		_ 1		HE9DG49
Gene No.	12	13	14	14	15	15	16

		ATCC Deposit		Z S S S S S S S S S S S S S S S S S S S	Total	5' NT 3' NT of of Clone Clone	3' NT of Clone	5° NT of Start	S' NT of AA I of AA of Do of AA of DO Start Signal NO:	¥8,0 Seo Seo Seo Seo Seo Seo Seo Seo Seo Seo	First AA of of Sig-		First AA of Secreted	Last AA of
Sene No.	cDNA Clone ID	Nr and Date	Vector	į×		-pocd-	.hac	Codon	Pep	¥	Pep	Pep	Portion	_
16	HE9DG49	97923 03/07/97 209071 05/22/97	Uni-ZAP XR	102	713	17	713	78	78	216	-	28	29	202
17	HELBA06	97923 03/07/97 209071	Uni-ZAP XR	- 27	1099	-	1099	38	38	141	1	22	23	215
17	HELBA06	97923 03/07/97 209071 05/27/97	Uni-ZAP XR	103	1080	_	1080	149	149	217	-	25	26	185
18	HSLFM29	97923 03/07/97 209071 05/22/97	Uni-ZAP XR	- 28	941	171	941	128	128	142		42	43	101-
19	HELBW38	97923 03/07/97 209071 05/22/97	Uni-ZAP XR	29	756	62	75.6	294	294	143	-	30	31	
20	HETHN28	97923 03/07/97 209071 05/22/97	Uni-ZAP XR	30	2100	408	2093	496	496	4 <u>4</u>	-	'		19

7	7
•	•

		<u>, , , , , , , , , , , , , , , , , , , </u>	<i>' '</i>				
Last AA of ORF	29	86	7	38	130	31	13
First AA of Secreted Portion		29		. 17	27	22	
Last AA of Sig Pep		28		16	26	21	
First of Sig Pep		-	1	ı.		-	
¥ãððö≻	145	146	147	148	149	150	151
5' NT of First AA of Signal Pep	567	21	210	242	178	44	1104
5' NT of Start Codon	267	21	210	242	178	4	1104
S' NT 3' NT of Of Clone Clone Seq. Seq.	1392	409	1322	710	1161	938	1581
5' NT of Clone Seq.	475	<u>'</u>		-	110	-	974
Total NT Seq.	1448	456	1326	710	1188	956	1603
× Š Đ Š K	31	-32	33	34	35	36	37
	ľ	Uni-ZAP XR	Uni-ZAP XR	Uni-ZAP XR	Uni-ZAP XR	Uni-ZAP XR	Uni-ZAP XR
ATCC Deposit Nr and Date	97923 03/07/97 209071 05/22/97	97923 03/07/97 209071 05/22/97	97923 03/07/97 209071 05/22/97	97923 03/07/97 209071 05/22/97	97923 03/07/97 209071 05/22/97	97923 03/07/97 209071 05/22/97	97924 03/07/97
cDNA Clone ID	<u>-</u> I			HFSBG13	HFTBE43		HKTAC77
Gene No.	21	777	23	24	25	26	27

							•							
Last AA of	ORF	-	/0	3	174	8	30	-68	\$, €	8 E	2/1	13/	47	4
First AA of Secreted		6.0	33		33	19	31	20	23	19	17	21	28	78
Last AA of Sig	Pep	C	35	Š	32	18	30	19	22	<u>∞</u>	87	20	-27	27
First AA of Sio	Pep	- -	-	1	I	I		-	-	-	<u> </u>	1		
¥ãa ġ	į, Y	751	5	134	155	156	157	158	218	159	<u> </u>	219	220	161
	7 T	209	119	180	126	43	171	55	28	. 17	5	72	54	269
5' NT of	Codon		119	581	126	43	1/1	55	28	17	15	72	54	269
3' NT of Clone	oed.	1067	629	1793	1123	875	843	489	489	534	1374	640	1399	965
		55	-	408	13		-	3	9	1	-	85	40	-
Total	Seq.	1089	629	1964	1522	875	843	489	489	534	1374	640	1529	969
FEGE	Ž×	38	39	9	41	42	43	4	104	45	46	105	106	47
	Vector	pBluescript	pBluescript	Lambda ZAP II	Uni-ZAP XR	Uni-ZAP XR	pSport1	Uni-ZAP XR	Uni-ZAP XR	Uni-ZAP XR	Uni-ZAP XR	Uni-ZAP XR	Uni-ZAP XR	Uni-ZAP XR
ATCC Deposit	Nr and Date	97924 03/07/97	97924 03/07/97	97924 03/07/97	97924	97924	97924	97924	97924	97924 03/07/97	97924 03/07/97	97924 03/07/97	97924 03/07/97	97924 03/07/97
	cDNA Clone ID	нгнѕнз6	96ASHTH	нговозе	HLTBX31	HLTCJ63	HMKAH44	HMQAJ64	HMQAJ64	HOABG65	HODCL36	HODCL36	HODCL36	HODCL50
	Gene No.	28	29	30	31	32	33	34	34	35	36	36	36	37

First AA Last of AA Secreted of Portion ORF	22	81	20 322	32 69	61 319	2282	30	17. 91	11 280	31 42	22	20 326	24 183
Last AA First of c Sig Seci Pep Por		17	19 2	31 3	9 09	21 2		18	30	30 3		19 2	23 2
First L AA A of of Sig Sig Pep P		-			-	-	-	<u> </u>	-	-1	_		1
¥ŠĐŠA ∀ŠĐŠA	162	_163	<u>1</u> 81	221	165	222	166	167	168	223	169	170	224
S' NT of First AA of Signal Pep	170	638	66	928	150	239	432	142	25	433	217	57	-35
5° NT of Start Codon	170	638	66	928	150	239	432	142	25	433	217	57	35
S' NT 3' NT of of Clone Clone Seq.	822	2020	2432	2435	2340	791	601	337	114F	1166	1148	809	586
5' NT of Clone Seq.	66	569	848	849	1627	92	188	-	_	21	63	164	4
Total NT Seq.	851	2020	2432	2435	2340	805	601	359	1141	1166	1560	1507	586
SEQ SEQ NO:	48	49	50	107	51	108	52	53	54	109	55	2 6	110
Vector	Uni-ZAP XR	Uni-ZAP XR	Uni-ZAP XR	Uni-ZAP XR	pBluescript SK-	pBluescript SK-	Uni-ZAP XR						
ATCC Deposit Nr and Date	97924 03/07/97	97924 03/07/97	97924 03/07/97	97924 03/07/97	97924 03/07/97	97924 03/07/97	97924 03/07/97	97924 03/07/97	97924 03/07/97	97924 03/07/97	97924 03/07/97	97924 03/07/97	97924 03/07/97
cDNA Clone ID	HODCV74	HODCZ16	HTOEU03	HTOEU03	HPBCJ74	HPBCJ74	HPMBU33	HSAUL66	HSIDQ18	HSIDQ18	HSJBB37	HSJBQ79	HSJBQ79
Gene No.	38	39	40	40	41	41	42	43	44	44	45	46	46

Gene cl No. Ck	47 HTI	48 HT	48 HT	49 HTI	50 HT	50 HT	51 HAI
cDNA Clone ID	HTEGA76	HTEJN13	HTEJN13	HTHBL86	HTSF071	HTSF071	HAPNO80
ATCC Deposit Nr and Date	97958 03/13/97 209072 05/22/97	97958 03/13/97 209072 05/22/97	97958 03/13/97 209072 05/22/97	97958 03/13/97 209072 05/22/97	97958 03/13/97 209072 05/22/97	97958 03/13/97 209072 05/22/97	209235 09/04/97
Vector	Uni-ZAP XR	Uni-ZAP XR	Uni-ZAP XR	Uni-ZAP XR	pBluescript	pBluescript	Uni-ZAP XR
XÖBÖ.	57	28		59	09	112	19
Total NT Seq.	450	1147	1134	777	1611	1333	1580
S' NT 3' NT of of Clone Clone Seq. Seq.	-	1			48	594	443
3' NT of Clone Seq.	450	1147	1134	777	598	1333	1554
S' NT of Start Codon	. 83	163	155	115	52	829	114
5' NT of First AA of Signal I	83	163	155	115	52		114
SEQ :	171	172	225	173	174	226	175
First AA of Sig Pep	·-	-	-	-			
Last AA I of Sig	35	15		<u>8</u>	30		_
First AA of Secreted Portion	36	16	20	19	31		2
Last AA of ORF	89	158	70	122	128	δ.	371

F A Et	137	N	-	Ţ.,	72	
AA AA ORF		215	54	22	102	47
First AA of Secreted Portion	29	29	- 33	21	34	39
Last AA of Sig Pep	78	28	32	20	33	38
First AA of Sig Pep		_	_	-	-	÷
¥ŠBŠ;≻	227	176	177	178	179	180
S' NT of AA I First SEQ AA of ID Signal NO: Pep Y	244	182	97	150	231	703
5' NT of Start Codon	244	182	97	150	231	703
S' NT 3' NT of of Clone Clone Seq. Seq.	708	1034	361	1638	1303	1011
5' NT of Clone Seq.	249	105_	-	-	35	655
Total NT Seq.	1015	1117	361	1668	1353	1011
Z B B S X	113	-62	63	2	65	99
Vector	Uni-ZAP XR	pBluescript	pSport1	Uni-ZAP XR	Uni-ZAP XR	Uni-ZAP XR
ATCC Deposit Nr and Date	97958 03/13/97 209072 05/22/97	97958 03/13/97 209072 05/22/97	97958 03/13/97 209072 05/22/97	97958 03/13/97 209072 05/22/97	97958 03/13/97 209072 05/22/97	97958 1 03/13/97 209072 05/22/97
	HAUCC47		HCFLD84	HE8EM69	HE8EZ48	HEBGF73
Gene No.	51	52	53	54	55	95

Last AA of ORF	95	8	26	- 10 -	64	21
First AA of Secreted Portion	36	30	22		20	22
Last AA of Sig Pep	35	29	21		61	21
First AA of Sig- Pep	1		1	<u> </u>	-	-
	181	182	183	184	185	186
5' NT of First AA of Signal	459		839	270	.272	127
5' NT of Start Codon	459	63	839	270	272	127
	1090	560	1581	711	935	484
S' NT 3' NT of of Clone Clone Seq.	267	-	765	∞ .	111	113
Fotal NT Seq.	1193	560	1657	711	935	504
× Še Še Še Še	19	89	69	70	71	72
Vector	Uni-ZAP XR	Uni-ZAP XR	pBluescript	Lambda ZAP II	Lambda ZAP II	Lambda ZAP II
ATCC Deposit Nr and Date	97958 03/13/97 209072 05/22/97	97958 03/13/97 209072	97958 03/13/97 209072 05/72/97	97958 03/13/97 209072	97958 03/13/97 209072 05/22/97	97958 03/13/97 209072 05/22/97
cDNA Clone ID	HFEBF41	HFRBU14	HFVGZ79	HHGCM76	нносов8	нндсР52
Gene No.	57	28	59	09	61	62

+ 17	· .		I		· · · · · ·	
Last AA of ORF	131	89	4	<u> </u>	22	169
First AA of Secreted Portion	19	33	28	37	12	15
Last AA of Sig Pep	- 18	32	27	36	11	14
First of Sig Pep	-	-		-	I	1
¥Š BŠ VÖÖ PŠ	187	188	189	190	228	192
5' NT of First AA of Signal Pep	96	248	630	167	575	187
5' NT of Start Codon	<u>,</u> 96	248		167		187
3' NT of Clone Seq.	620	581	1786	008	1076	1888
S' NT 3' NT of of Clone Clone Seq.		156	537	116	398	18
Total NT Seq.	620	581	1843	1441	1076	2776
X SEQ	73	74	75	92	114	78
Vector	Lambda ZAP II	Lambda ZAP II	Uni-ZAP XR	Uni-ZAP XR	Uni-ZAP XR	pCMVSport 3.0
ATCC Deposit Nr and Date	97958 03/13/97 209072 05/22/97	97958 03/13/97 209072 05/22/97	97958 03/13/97 209072 05/22/97	97958 03/13/97 209072 05/22/97	97958 03/13/97 209072 05/22/97	97958 03/13/97 209072 05/22/97
cDNA Clone ID	HHGDB72	HHGDI71	HHSD145	HHSEB66	HJPAZ83	HLDBO49
Gene No.	63	4 9	65	99	<i>L</i> 9	89

			, 				
Last AA of ORF	92	131	16	175	69	24	72
First AA of Secreted Portion	23	23	33	- 24	27_	21	56
Last AA of Sig Pep	22	22	32	23	26	20	25
First AA of Sig Pep	-	I	1	-	- ·	_	-
	193	229	-194	195	196	197	198
5' NT of First SEQ AA of ID Signal NO: Pep Y	534	534	40	238	286	58	14
5' NT of Start Codon	534	534	40	238	286	- 58	14
3' NT of Clone Seq.	1480	1487	1077	780	770	481	623
5' NT 3' NT of Clone Clone Seq. Seq.	401	401	33	18	101	-	 '
Total NT Seq.	1525	1487	1563	1020	0//	481	4 8
NT SEQ NO:	6/	115	08	8	28	83	84
Vector	pCMVSport 3.0		Uni-ZAP XR		Uni-ZAP XR	Uni-ZAP XR	pCMVSport 3.0
ATCC Deposit Nr and Date	97958 03/13/97 209072 05/22/97	209226 08/28/97	97958 03/13/97 209072 05/22/97	97957 03/13/97 209073 05/22/97	97957 03/13/97 209073 05/22/97	97957 03/13/97 209073 05/22/97	97957 03/13/97 209073 05/22/97
cDNA Clone ID	HLDBQ19		l .		ŀ	1	HNTAC73
Gene No.	69	69	70	71	72	73	74

Last Of ORF	288	27	623	09	648	28
First AA of Secreted Portion	13		31	33	31	22
First Last AA AA of of Sig Pep Pep	12		30	32	99	21.
First AA of Sig Pep	1	-	-	-	-	
¥8. BSBS ×	199	230	200	231	201	232
S' NT of First SEC AA of ID Signal NO: Pep Y	86	545	56	477	251	677
5' NT of Start Codon	86		56	477	251	677
S' NT 3' NT of of Clone Clone Seq. Seq.	1284	1283	1747	1747	2566	1098
S' NT 3' NT of Clone Clone Seq. Seq.	435	428	290	288	1843 2566	375
Total NT Seq.	1351	1350	2527	2527	2566	1098
SEQ NO:	-88	116	98	1:17	87	118
Vector	Uni-ZAP XR					
ATCC Deposit Nr and Date	97957 03/13/97 209073 05/22/97	97957 03/13/97 209073 05/22/97	97957 03/13/97 209073 05/22/97	97957 03/13/97 209073 05/22/97	97957 03/13/97 209073 05/22/97	97957 03/13/97 209073 05/22/97
cDNA Clone ID	HOSEI45	HOSEI45	HOSFD58	HOSFD58	HSAUM95	HSAUM95
Gene No.	75	75	92	92	11	77

Last AA of ORF	54	265	17	314	206	194
First AA of Secreted Portion	33	12		50	21	70
Last AA of Sig Pep	32	=	,	61	70	69
irst AA of Sig Pep	1	I	_	T .		-
AA SEQ D NO: Y	202	203	233	204	205	206
Start Signal NO: Codon Pep Y 1	83	188	315	92	414	157
\$	83	188	315	92	414	157
3' NT of Clone Seq.	540	1165	1166	2449	2058	1411
5' NT 3' NT of of Clone Clone Seq.		152	152	1149	476	345
Total NT Seq.	540	1863	1679	2478	2058	1411
X Ö B Ö Z	80 80	68	119	8	· 16	92
Vector	Uni-ZAP XR	pBluescript				
ATCC Deposit Nr and Date	97957 03/13/97 209073 05/22/97	97957 03/13/97 209073 05/22/97	97957 03/13/97 209073 05/22/97	97957 03/13/97 209073 05/22/97	97957 03/13/97 209073 05/22/97	97957 03/13/97 209073 05/22/97
cDNA Clone ID	HSAUR67	HSKD181	HSKD181	HSKDW91	HTLEX50	HSKHL65
Gene No.	78	79	79	80	81	82

Last AA of ORF	71	329	95	57	391	25
First AA of Secreted Portion (38	31	50	21	7	52
Last AA of Sig Pep	37	30	61	70		21
		-	-		'	-
SEQ H	235	207	236	208	209	210
5' N7 of First AA o Signa Pep	526	397	228	445	523	117
5' NT of Start Codon	526	397	228	445	523	117
S' NT 3' NT of of Clone Clone Seq. Seq.	1411	147- 2184	2063	809	2394	672
S' NT 3' NT of of Clone Clone Seq. Seq.	345	147-	138	524	481	-
Total NT Seq.	1411	2187	2256	757	2394	672
Z S O S ×	121	-93	122	46	95	96
Vector	pBluescript	Uni-ZAP XR	Uni-ZAP XR	Uni-ZAP XR	ZAP Express	Uni-ZAP XR
ATCC Deposit Nr and Date	97957 03/13/97 209073 05/22/97	97957 03/13/97 209073 05/22/97	97957 03/13/97 209073 05/22/97	97957 03/13/97 209073 05/22/97	97957 03/13/97 209073 05/22/97	97957 03/13/97 209073 05/22/97
cDNA Clone ID	HSKHL65	HHFGA11	411	HWTBL40	l_	HCACY32
Gene No.	82	83	83	84	85	98

of AA First Last AA First AA Last AA of D of of of of AA Screeted of AA Pep Pep Portion ORF	37
Hirst A First A Secre	21
AAA Of Sign	20
Firs Of Of Sig Per	-
ŞEŞ ŞÜĞĞĞ Y	211
5' NT of First AA of Signal Pep	207 211
5' NT 3' NT of of 5' NT F of of Clone Clone of A Clone Seq. Seq. Start Sign.	207
3' NT of Clone Seq.	1419
5' NT of Clone Seq.	-
Se Z d	1419
X Š B Š K	25
Vector	Uni-ZAP XR 97
ATCC Deposit Nr and Date	97957 03/13/97 209073 05/22/97
Gene CDNA Clone ID	HCED021
ene Vo.	87

10

15

20

25

Table 1 summarizes the information corresponding to each "Gene No." described above. The nucleotide sequence identified as "NT SEQ ID NO:X" was assembled from partially homologous ("overlapping") sequences obtained from the "cDNA clone ID" identified in Table 1 and, in some cases, from additional related DNA clones. The overlapping sequences were assembled into a single contiguous sequence of high redundancy (usually three to five overlapping sequences at each nucleotide position), resulting in a final sequence identified as SEQ ID NO:X.

The cDNA Clone ID was deposited on the date and given the corresponding deposit number listed in "ATCC Deposit No:Z and Date." Some of the deposits contain multiple different clones corresponding to the same gene. "Vector" refers to the type of vector contained in the cDNA Clone ID.

"Total NT Seq." refers to the total number of nucleotides in the contig identified by "Gene No." The deposited clone may contain all or most of these sequences, reflected by the nucleotide position indicated as "5' NT of Clone Seq." and the "3' NT of Clone Seq." of SEQ ID NO:X. The nucleotide position of SEQ ID NO:X of the putative start codon (methionine) is identified as "5' NT of Start Codon." Similarly, the nucleotide position of SEQ ID NO:X of the predicted signal sequence is identified as "5' NT of First AA of Signal Pep."

The translated amino acid sequence, beginning with the methionine, is identified as "AA SEQ ID NO:Y," although other reading frames can also be easily translated using known molecular biology techniques. The polypeptides produced by these alternative open reading frames are specifically contemplated by the present invention.

The first and last amino acid position of SEQ ID NO:Y of the predicted signal peptide is identified as "First AA of Sig Pep" and "Last AA of Sig Pep." The predicted first amino acid position of SEQ ID NO:Y of the secreted portion is identified as "Predicted First AA of Secreted Portion." Finally, the amino acid position of SEQ ID NO:Y of the last amino acid in the open reading frame is identified as "Last AA of ORF."

SEQ ID NO:X and the translated SEQ ID NO:Y are sufficiently accurate and otherwise suitable for a variety of uses well known in the art and described further below. For instance, SEQ ID NO:X is useful for designing nucleic acid hybridization probes that will detect nucleic acid sequences contained in SEQ ID NO:X or the cDNA contained in the deposited clone. These probes will also hybridize to nucleic acid molecules in biological samples, thereby enabling a variety of forensic and diagnostic methods of the invention. Similarly, polypeptides identified from SEQ ID NO:Y may be used to generate antibodies which bind specifically to the secreted proteins encoded by the cDNA clones identified in Table 1.

10

15

20

25

30

35

Nevertheless, DNA sequences generated by sequencing reactions can contain sequencing errors. The errors exist as misidentified nucleotides, or as insertions or deletions of nucleotides in the generated DNA sequence. The erroneously inserted or deleted nucleotides cause frame shifts in the reading frames of the predicted amino acid sequence. In these cases, the predicted amino acid sequence diverges from the actual amino acid sequence, even though the generated DNA sequence may be greater than 99.9% identical to the actual DNA sequence (for example, one base insertion or deletion in an open reading frame of over 1000 bases).

Accordingly, for those applications requiring precision in the nucleotide sequence or the amino acid sequence, the present invention provides not only the generated nucleotide sequence identified as SEQ ID NO:X and the predicted translated amino acid sequence identified as SEQ ID NO:Y, but also a sample of plasmid DNA containing a human cDNA of the invention deposited with the ATCC, as set forth in Table 1. The nucleotide sequence of each deposited clone can readily be determined by sequencing the deposited clone in accordance with known methods. The predicted amino acid sequence can then be verified from such deposits. Moreover, the amino acid sequence of the protein encoded by a particular clone can also be directly determined by peptide sequencing or by expressing the protein in a suitable host cell containing the deposited human cDNA, collecting the protein, and determining its sequence.

The present invention also relates to the genes corresponding to SEQ ID NO:X, SEQ ID NO:Y, or the deposited clone. The corresponding gene can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include preparing probes or primers from the disclosed sequence and identifying or amplifying the corresponding gene from appropriate sources of genomic material.

Also provided in the present invention are species homologs. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source for the desired homologue.

The polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

The polypeptides may be in the form of the secreted protein, including the mature form, or may be a part of a larger protein, such as a fusion protein (see below).

10

15

20

25

30

35

It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification, such as multiple histidine residues, or an additional sequence for stability during recombinant production.

The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced version of a polypeptide, including the secreted polypeptide, can be substantially purified by the one-step method described in Smith and Johnson, Gene 67:31-40 (1988). Polypeptides of the invention also can be purified from natural or recombinant sources using antibodies of the invention raised against the secreted protein in methods which are well known in the art.

Signal Sequences

Methods for predicting whether a protein has a signal sequence, as well as the cleavage point for that sequence, are available. For instance, the method of McGeoch, Virus Res. 3:271-286 (1985), uses the information from a short N-terminal charged region and a subsequent uncharged region of the complete (uncleaved) protein. The method of von Heinje, Nucleic Acids Res. 14:4683-4690 (1986) uses the information from the residues surrounding the cleavage site, typically residues -13 to +2, where +1 indicates the amino terminus of the secreted protein. The accuracy of predicting the cleavage points of known mammalian secretory proteins for each of these methods is in the range of 75-80%. (von Heinje, supra.) However, the two methods do not always produce the same predicted cleavage point(s) for a given protein.

In the present case, the deduced amino acid sequence of the secreted polypeptide was analyzed by a computer program called SignalP (Henrik Nielsen et al., Protein Engineering 10:1-6 (1997)), which predicts the cellular location of a protein based on the amino acid sequence. As part of this computational prediction of localization, the methods of McGeoch and von Heinje are incorporated. The analysis of the amino acid sequences of the secreted proteins described herein by this program provided the results shown in Table 1.

As one of ordinary skill would appreciate, however, cleavage sites sometimes vary from organism to organism and cannot be predicted with absolute certainty. Accordingly, the present invention provides secreted polypeptides having a sequence shown in SEQ ID NO:Y which have an N-terminus beginning within 5 residues (i.e., + or - 5 residues) of the predicted cleavage point. Similarly, it is also recognized that in some cases, cleavage of the signal sequence from a secreted protein is not entirely

15

20

25

30

35

uniform, resulting in more than one secreted species. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

Moreover, the signal sequence identified by the above analysis may not necessarily predict the naturally occurring signal sequence. For example, the naturally occurring signal sequence may be further upstream from the predicted signal sequence. However, it is likely that the predicted signal sequence will be capable of directing the secreted protein to the ER. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

10 Polynucleotide and Polypeptide Variants

"Variant" refers to a polynucleotide or polypeptide differing from the polynucleotide or polypeptide of the present invention, but retaining essential properties thereof. Generally, variants are overall closely similar, and, in many regions, identical to the polynucleotide or polypeptide of the present invention.

By a polynucleotide having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence of the present invention, it is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the polypeptide. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. The query sequence may be an entire sequence shown in Table 1, the ORF (open reading frame), or any fragement specified as described herein.

As a practical matter, whether any particular nucleic acid molecule or polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to a nucleotide sequence of the presence invention can be determined conventionally using known computer programs. A preferred method for determing the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are both DNA sequences. An RNA sequence can be compared by converting U's to T's. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB alignment of DNA sequences to calculate percent identity are:

Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization

10

15

20

25

30

35

Group Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty 0.05, Window Size=500 or the length of the subject nucleotide sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence because of 5' or 3' deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for 5' and 3' truncations of the subject sequence when calculating percent identity. For subject sequences truncated at the 5' or 3' ends, relative to the the query sequence, the percent identity is corrected by calculating the number of bases of the query sequence that are 5' and 3' of the subject sequence, which are not matched/aligned, as a percent of the total bases of the query sequence. Whether a nucleotide is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This corrected score is what is used for the purposes of the present invention. Only bases outside the 5' and 3' bases of the subject sequence, as displayed by the FASTDB alignment, which are not matched/aligned with the query sequence, are calculated for the purposes of manually adjusting the percent identity score.

For example, a 90 base subject sequence is aligned to a 100 base query sequence to determine percent identity. The deletions occur at the 5' end of the subject sequence and therefore, the FASTDB alignment does not show a matched/alignement of the first 10 bases at 5' end. The 10 unpaired bases represent 10% of the sequence (number of bases at the 5' and 3' ends not matched/total number of bases in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 bases were perfectly matched the final percent identity would be 90%. In another example, a 90 base subject sequence is compared with a 100 base query sequence. This time the deletions are internal deletions so that there are no bases on the 5' or 3' of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only bases 5' and 3' of the subject sequence which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are to made for the purposes of the present invention.

By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to a query amino acid sequence of the present invention, it is intended that the amino acid sequence of the subject polypeptide is identical to the query sequence except that the subject polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the query amino acid sequence. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a query

10

15

20

amino acid sequence, up to 5% of the amino acid residues in the subject sequence may be inserted, deleted, (indels) or substituted with another amino acid. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the amino acid sequences shown in Table 1 or to the amino acid sequence encoded by deposited DNA clone can be determined conventionally using known computer programs. A preferred method for determing the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are either both nucleotide sequences or both amino acid sequences. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB amino acid alignment are: Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject amino acid sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence due to N- or Cterminal deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for N- and Cterminal truncations of the subject sequence when calculating global percent identity. For subject sequences truncated at the N- and C-termini, relative to the the query 25 sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding subject residue, as a percent of the total bases of the query sequence. Whether a residue is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent 30 identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what is used for the purposes of the present invention. Only residues to the N- and C-termini of the subject sequence, which are not matched/aligned with the query sequence, are considered for the purposes of manually adjusting the percent identity score. That is, 35 only query residue positions outside the farthest N- and C-terminal residues of the subject sequence.

For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the Nterminus of the subject sequence and therefore, the FASTDB alignment does not show a matching/alignment of the first 10 residues at the N-terminus. The 10 unpaired residues represent 10% of the sequence (number of residues at the N- and C- termini 5 not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 residues were perfectly matched the final percent identity would be 90%. In another example, a 90 residue subject sequence is compared with a 100 residue query sequence. This time the deletions are internal deletions so there are no residues at the N- or C-10 termini of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject sequence, as displayed in the FASTDB alignment, which are not matched/aligned with the query sequnce are manually corrected for. No other manual corrections are to made for the 15 purposes of the present invention.

The variants may contain alterations in the coding regions, non-coding regions, or both. Especially preferred are polynucleotide variants containing alterations which produce silent substitutions, additions, or deletions, but do not alter the properties or activities of the encoded polypeptide. Nucleotide variants produced by silent substitutions due to the degeneracy of the genetic code are preferred. Moreover, variants in which 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination are also preferred. Polynucleotide variants can be produced for a variety of reasons, e.g., to optimize codon expression for a particular host (change codons in the human mRNA to those preferred by a bacterial host such as E. coli).

Naturally occurring variants are called "allelic variants," and refer to one of several alternate forms of a gene occupying a given locus on a chromosome of an organism. (Genes II, Lewin, B., ed., John Wiley & Sons, New York (1985).) These allelic variants can vary at either the polynucleotide and/or polypeptide level. Alternatively, non-naturally occurring variants may be produced by mutagenesis techniques or by direct synthesis.

Using known methods of protein engineering and recombinant DNA technology, variants may be generated to improve or alter the characteristics of the polypeptides of the present invention. For instance, one or more amino acids can be deleted from the N-terminus or C-terminus of the secreted protein without substantial loss of biological function. The authors of Ron et al., J. Biol. Chem. 268: 2984-2988 (1993), reported variant KGF proteins having heparin binding activity even after

20

25

30

10

15

20

25

30

35

deleting 3, 8, or 27 amino-terminal amino acid residues. Similarly, Interferon gamma exhibited up to ten times higher activity after deleting 8-10 amino acid residues from the carboxy terminus of this protein. (Dobeli et al., J. Biotechnology 7:199-216 (1988).)

Moreover, ample evidence demonstrates that variants often retain a biological activity similar to that of the naturally occurring protein. For example, Gayle and coworkers (J. Biol. Chem 268:22105-22111 (1993)) conducted extensive mutational analysis of human cytokine IL-1a. They used random mutagenesis to generate over 3,500 individual IL-1a mutants that averaged 2.5 amino acid changes per variant over the entire length of the molecule. Multiple mutations were examined at every possible amino acid position. The investigators found that "[m]ost of the molecule could be altered with little effect on either [binding or biological activity]." (See, Abstract.) In fact, only 23 unique amino acid sequences, out of more than 3,500 nucleotide sequences examined, produced a protein that significantly differed in activity from wild-type.

Furthermore, even if deleting one or more amino acids from the N-terminus or C-terminus of a polypeptide results in modification or loss of one or more biological functions, other biological activities may still be retained. For example, the ability of a deletion variant to induce and/or to bind antibodies which recognize the secreted form will likely be retained when less than the majority of the residues of the secreted form are removed from the N-terminus or C-terminus. Whether a particular polypeptide lacking N- or C-terminal residues of a protein retains such immunogenic activities can readily be determined by routine methods described herein and otherwise known in the art.

Thus, the invention further includes polypeptide variants which show substantial biological activity. Such variants include deletions, insertions, inversions, repeats, and substitutions selected according to general rules known in the art so as have little effect on activity. For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie, J. U. et al., Science 247:1306-1310 (1990), wherein the authors indicate that there are two main strategies for studying the tolerance of an amino acid sequence to change.

The first strategy exploits the tolerance of amino acid substitutions by natural selection during the process of evolution. By comparing amino acid sequences in different species, conserved amino acids can be identified. These conserved amino acids are likely important for protein function. In contrast, the amino acid positions where substitutions have been tolerated by natural selection indicates that these positions are not critical for protein function. Thus, positions tolerating amino acid substitution could be modified while still maintaining biological activity of the protein.

WO 98/42738 PCT/US98/05311

97

The second strategy uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene to identify regions critical for protein function. For example, site directed mutagenesis or alanine-scanning mutagenesis (introduction of single alanine mutations at every residue in the molecule) can be used. (Cunningham and Wells, Science 244:1081-1085 (1989).) The resulting mutant molecules can then be tested for biological activity.

As the authors state, these two strategies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes are likely to be permissive at certain amino acid positions in the protein. For example, most buried (within the tertiary structure of the protein) amino—acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved. Moreover, tolerated conservative amino acid substitutions involve replacement of the aliphatic or hydrophobic amino acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic residues Asp and Glu; replacement of the amide residues Asn and Gln, replacement of the basic residues Lys, Arg, and His; replacement of the aromatic residues Phe, Tyr, and Trp, and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly.

Besides conservative amino acid substitution, variants of the present invention include (i) substitutions with one or more of the non-conserved amino acid residues, where the substituted amino acid residues may or may not be one encoded by the genetic code, or (ii) substitution with one or more of amino acid residues having a substituent group, or (iii) fusion of the mature polypeptide with another compound, such as a compound to increase the stability and/or solubility of the polypeptide (for example, polyethylene glycol), or (iv) fusion of the polypeptide with additional amino acids, such as an IgG Fc fusion region peptide, or leader or secretory sequence, or a sequence facilitating purification. Such variant polypeptides are deemed to be within the scope of those skilled in the art from the teachings herein.

For example, polypeptide variants containing amino acid substitutions of charged amino acids with other charged or neutral amino acids may produce proteins with improved characteristics, such as less aggregation. Aggregation of pharmaceutical formulations both reduces activity and increases clearance due to the aggregate's immunogenic activity. (Pinckard et al., Clin. Exp. Immunol. 2:331-340 (1967); Robbins et al., Diabetes 36: 838-845 (1987); Cleland et al., Crit. Rev. Therapeutic Drug Carrier Systems 10:307-377 (1993).)

30

5

10

15

20

10

15

20

25

30

35

Polynucleotide and Polypeptide Fragments

In the present invention, a "polynucleotide fragment" refers to a short polynucleotide having a nucleic acid sequence contained in the deposited clone or shown in SEQ ID NO:X. The short nucleotide fragments are preferably at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably, at least about 40 nt in length. A fragment "at least 20 nt in length," for example, is intended to include 20 or more contiguous bases from the cDNA sequence contained in the deposited clone or the nucleotide sequence shown in SEQ ID NO:X. These nucleotide fragments are useful as diagnostic probes and primers as discussed herein. Of course, larger fragments (e.g., 50, 150, 500, 600, 2000 nucleotides) are preferred.

Moreover, representative examples of polynucleotide fragments of the invention, include, for example, fragments having a sequence from about nucleotide number 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, 451-500, 501-550, 551-600, 651-700, or 701 to the end of SEQ ID NO:X or the cDNA contained in the deposited clone. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini. Preferably, these fragments encode a polypeptide which has biological activity.

In the present invention, a "polypeptide fragment" refers to a short amino acid sequence contained in SEQ ID NO:Y or encoded by the cDNA contained in the deposited clone. Protein fragments may be "free-standing," or comprised within a larger polypeptide of which the fragment forms a part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments from about amino acid number 1-20, 21-40, 41-60, 61-80, 81-100, 102-120, 121-140, 141-160, or 161 to the end of the coding region. Moreover, polypeptide fragments can be about 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, or 150 amino acids in length. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) amino acids, at either extreme or at both extremes.

Preferred polypeptide fragments include the secreted protein as well as the mature form. Further preferred polypeptide fragments include the secreted protein or the mature form having a continuous series of deleted residues from the amino or the carboxy terminus, or both. For example, any number of amino acids, ranging from 1-60, can be deleted from the amino terminus of either the secreted polypeptide or the mature form. Similarly, any number of amino acids, ranging from 1-30, can be deleted from the carboxy terminus of the secreted protein or mature form. Furthermore, any

10

15

20

25

30

35

combination of the above amino and carboxy terminus deletions are preferred. Similarly, polynucleotide fragments encoding these polypeptide fragments are also preferred.

Also preferred are polypeptide and polynucleotide fragments characterized by structural or functional domains, such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions. Polypeptide fragments of SEQ ID NO:Y falling within conserved domains are specifically contemplated by the present invention. Moreover, polynucleotide fragments encoding these domains are also contemplated.

Other preferred fragments are biologically active fragments. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide of the present invention. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity.

Epitopes & Antibodies

In the present invention, "epitopes" refer to polypeptide fragments having antigenic or immunogenic activity in an animal, especially in a human. A preferred embodiment of the present invention relates to a polypeptide fragment comprising an epitope, as well as the polynucleotide encoding this fragment. A region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." In contrast, an "immunogenic epitope" is defined as a part of a protein that elicits an antibody response. (See, for instance, Geysen et al., Proc. Natl. Acad. Sci. USA 81:3998- 4002 (1983).)

Fragments which function as epitopes may be produced by any conventional means. (See, e.g., Houghten, R. A., Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985) further described in U.S. Patent No. 4,631,211.)

In the present invention, antigenic epitopes preferably contain a sequence of at least seven, more preferably at least nine, and most preferably between about 15 to about 30 amino acids. Antigenic epitopes are useful to raise antibodies, including monoclonal antibodies, that specifically bind the epitope. (See, for instance, Wilson et al., Cell 37:767-778 (1984); Sutcliffe, J. G. et al., Science 219:660-666 (1983).)

Similarly, immunogenic epitopes can be used to induce antibodies according to methods well known in the art. (See, for instance, Sutcliffe et al., supra; Wilson et al., supra; Chow, M. et al., Proc. Natl. Acad. Sci. USA 82:910-914; and Bittle, F. J. et

10

20

25

30

35

al., J. Gen. Virol. 66:2347-2354 (1985).) A preferred immunogenic epitope includes the secreted protein. The immunogenic epitopes may be presented together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse) or, if it is long enough (at least about 25 amino acids), without a carrier. However, immunogenic epitopes comprising as few as 8 to 10 amino acids have been shown to be sufficient to raise antibodies capable of binding to, at the very least, linear epitopes in a denatured polypeptide (e.g., in Western blotting.)

As used herein, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab')2 fragments) which are capable of specifically binding to protein. Fab and F(ab')2 fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding than an intact antibody. (Wahl et al., J. Nucl. Med. 24:316-325 (1983).) Thus, these fragments are preferred, as well as the products of a FAB or other immunoglobulin expression library. Moreover, antibodies of the present invention include chimeric, single chain, and

15 humanized antibodies.

Fusion Proteins

Any polypeptide of the present invention can be used to generate fusion proteins. For example, the polypeptide of the present invention, when fused to a second protein, can be used as an antigenic tag. Antibodies raised against the polypeptide of the present invention can be used to indirectly detect the second protein by binding to the polypeptide. Moreover, because secreted proteins target cellular locations based on trafficking signals, the polypeptides of the present invention can be used as targeting molecules once fused to other proteins.

Examples of domains that can be fused to polypeptides of the present invention include not only heterologous signal sequences, but also other heterologous functional regions. The fusion does not necessarily need to be direct, but may occur through linker sequences.

Moreover, fusion proteins may also be engineered to improve characteristics of the polypeptide of the present invention. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence during purification from the host cell or subsequent handling and storage. Also, peptide moieties may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to facilitate handling of polypeptides are familiar and routine techniques in the art.

WO 98/42738 PCT/US98/05311

101

Moreover, polypeptides of the present invention, including fragments, and specifically epitopes, can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. One reported example describes chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP A 394,827; Traunecker et al., Nature 331:84-86 (1988).) Fusion proteins having disulfide-linked dimeric structures (due to the IgG) can also be more efficient in binding and neutralizing other molecules, than the monomeric secreted protein or protein fragment alone. (Fountoulakis et al., J. Biochem. 270:3958-3964 (1995).)

Similarly, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP-A 0232 262.) Alternatively, deleting the Fc part after the fusion protein has been expressed, detected, and purified, would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. (See, D. Bennett et al., J. Molecular Recognition 8:52-58 (1995); K. Johanson et al., J. Biol. Chem. 270:9459-9471 (1995).)

Moreover, the polypeptides of the present invention can be fused to marker sequences, such as a peptide which facilitates purification of the fused polypeptide. In preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci. USA 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. Another peptide tag useful for purification, the "HA" tag, corresponds to an epitope derived from the influenza hemagglutinin protein. (Wilson et al., Cell 37:767 (1984).)

Thus, any of these above fusions can be engineered using the polynucleotides or the polypeptides of the present invention.

35

30

5

10

15

20

10

15

20

25

30

35

Vectors, Host Cells, and Protein Production

The present invention also relates to vectors containing the polynucleotide of the present invention, host cells, and the production of polypeptides by recombinant techniques. The vector may be, for example, a phage, plasmid, viral, or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

The polynucleotide insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the E. coli lac, trp, phoA and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance genes for culturing in E. coli and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as E. coli, Streptomyces and Salmonella typhimurium cells; fungal cells, such as yeast cells; insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS, 293, and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from QIAGEN, Inc.; pBluescript vectors, Phagescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene Cloning Systems, Inc.; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia Biotech, Inc. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

10

15

20

25

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., Basic Methods In Molecular Biology (1986). It is specifically contemplated that the polypeptides of the present invention may in fact be expressed by a host cell lacking a recombinant vector.

A polypeptide of this invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Polypeptides of the present invention, and preferably the secreted form, can also be recovered from: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect, and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins, this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

30 Uses of the Polynucleotides

Each of the polynucleotides identified herein can be used in numerous ways as reagents. The following description should be considered exemplary and utilizes known techniques.

The polynucleotides of the present invention are useful for chromosome identification. There exists an ongoing need to identify new chromosome markers, since few chromosome marking reagents, based on actual sequence data (repeat

10

15

20

25

30

35

polymorphisms), are presently available. Each polynucleotide of the present invention can be used as a chromosome marker.

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the sequences shown in SEQ ID NO:X. Primers can be selected using computer analysis so that primers do not span more than one predicted exon in the genomic DNA. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the SEQ ID NO:X will yield an amplified fragment.

Similarly, somatic hybrids provide a rapid method of PCR mapping the polynucleotides to particular chromosomes. Three or more clones can be assigned per day using a single thermal cycler. Moreover, sublocalization of the polynucleotides can be achieved with panels of specific chromosome fragments. Other gene mapping strategies that can be used include in situ hybridization, prescreening with labeled flow-sorted chromosomes, and preselection by hybridization to construct chromosome specific-cDNA libraries.

Precise chromosomal location of the polynucleotides can also be achieved using fluorescence in situ hybridization (FISH) of a metaphase chromosomal spread. This technique uses polynucleotides as short as 500 or 600 bases; however, polynucleotides 2,000-4,000 bp are preferred. For a review of this technique, see Verma et al., "Human Chromosomes: a Manual of Basic Techniques," Pergamon Press, New York (1988).

For chromosome mapping, the polynucleotides can be used individually (to mark a single chromosome or a single site on that chromosome) or in panels (for marking multiple sites and/or multiple chromosomes). Preferred polynucleotides correspond to the noncoding regions of the cDNAs because the coding sequences are more likely conserved within gene families, thus increasing the chance of cross hybridization during chromosomal mapping.

Once a polynucleotide has been mapped to a precise chromosomal location, the physical position of the polynucleotide can be used in linkage analysis. Linkage analysis establishes coinheritance between a chromosomal location and presentation of a particular disease. (Disease mapping data are found, for example, in V. McKusick, Mendelian Inheritance in Man (available on line through Johns Hopkins University Welch Medical Library).) Assuming 1 megabase mapping resolution and one gene per 20 kb, a cDNA precisely localized to a chromosomal region associated with the disease could be one of 50-500 potential causative genes.

Thus, once coinheritance is established, differences in the polynucleotide and the corresponding gene between affected and unaffected individuals can be examined.

10

15

20

25

30

35

First, visible structural alterations in the chromosomes, such as deletions or translocations, are examined in chromosome spreads or by PCR. If no structural alterations exist, the presence of point mutations are ascertained. Mutations observed in some or all affected individuals, but not in normal individuals, indicates that the mutation may cause the disease. However, complete sequencing of the polypeptide and the corresponding gene from several normal individuals is required to distinguish the mutation from a polymorphism. If a new polymorphism is identified, this polymorphic polypeptide can be used for further linkage analysis.

Furthermore, increased or decreased expression of the gene in affected individuals as compared to unaffected individuals can be assessed using polynucleotides of the present invention. Any of these alterations (altered expression, chromosomal rearrangement, or mutation) can be used as a diagnostic or prognostic marker.

In addition to the foregoing, a polynucleotide can be used to control gene expression through triple helix formation or antisense DNA or RNA. Both methods rely on binding of the polynucleotide to DNA or RNA. For these techniques, preferred polynucleotides are usually 20 to 40 bases in length and complementary to either the region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxy-nucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988).) Triple helix formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques are effective in model systems, and the information disclosed herein can be used to design antisense or triple helix polynucleotides in an effort to treat disease.

Polynucleotides of the present invention are also useful in gene therapy. One goal of gene therapy is to insert a normal gene into an organism having a defective gene, in an effort to correct the genetic defect. The polynucleotides disclosed in the present invention offer a means of targeting such genetic defects in a highly accurate manner. Another goal is to insert a new gene that was not present in the host genome, thereby producing a new trait in the host cell.

The polynucleotides are also useful for identifying individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identifying

personnel. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The polynucleotides of the present invention can be used as additional DNA markers for RFLP.

5

10

15

20

25

The polynucleotides of the present invention can also be used as an alternative to RFLP, by determining the actual base-by-base DNA sequence of selected portions of an individual's genome. These sequences can be used to prepare PCR primers for amplifying and isolating such selected DNA, which can then be sequenced. Using this technique, individuals can be identified because each individual will have a unique set of DNA sequences. Once an unique ID database is established for an individual, positive identification of that individual, living or dead, can be made from extremely small tissue samples.

Forensic biology also benefits from using DNA-based identification techniques as disclosed herein. DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, semen, etc., can be amplified using PCR. In one prior art technique, gene sequences amplified from polymorphic loci, such as DQa class II HLA gene, are used in forensic biology to identify individuals. (Erlich, H., PCR Technology, Freeman and Co. (1992).) Once these specific polymorphic loci are amplified, they are digested with one or more restriction enzymes, yielding an identifying set of bands on a Southern blot probed with DNA corresponding to the DQa class II HLA gene. Similarly, polynucleotides of the present invention can be used as polymorphic markers for forensic purposes.

There is also a need for reagents capable of identifying the source of a particular tissue. Such need arises, for example, in forensics when presented with tissue of unknown origin. Appropriate reagents can comprise, for example, DNA probes or primers specific to particular tissue prepared from the sequences of the present invention. Panels of such reagents can identify tissue by species and/or by organ type. In a similar fashion, these reagents can be used to screen tissue cultures for contamination.

30

35

In the very least, the polynucleotides of the present invention can be used as molecular weight markers on Southern gels, as diagnostic probes for the presence of a specific mRNA in a particular cell type, as a probe to "subtract-out" known sequences in the process of discovering novel polynucleotides, for selecting and making oligomers for attachment to a "gene chip" or other support, to raise anti-DNA antibodies using DNA immunization techniques, and as an antigen to elicit an immune response.

10

15

20

25

30

35

Uses of the Polypeptides

Each of the polypeptides identified herein can be used in numerous ways. The following description should be considered exemplary and utilizes known techniques.

A polypeptide of the present invention can be used to assay protein levels in a biological sample using antibody-based techniques. For example, protein expression in tissues can be studied with classical immunohistological methods. (Jalkanen, M., et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, M., et al., J. Cell. Biol. 105:3087-3096 (1987).) Other antibody-based methods useful for detecting protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine (125I, 121I), carbon (14C), sulfur (35S), tritium (3H), indium (112In), and technetium (99mTc), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

In addition to assaying secreted protein levels in a biological sample, proteins can also be detected in vivo by imaging. Antibody labels or markers for in vivo imaging of protein include those detectable by X-radiography, NMR or ESR. For X-radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma.

A protein-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example, 131I, 112In, 99mTc), a radio-opaque substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously, or intraperitoneally) into the mammal. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of 99mTc. The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain the specific protein. In vivo tumor imaging is described in S.W. Burchiel et al., "Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments." (Chapter 13 in Tumor Imaging: The Radiochemical Detection of Cancer, S.W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982).)

10

15

20

25

30

Thus, the invention provides a diagnostic method of a disorder, which involves (a) assaying the expression of a polypeptide of the present invention in cells or body fluid of an individual; (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of a disorder.

Moreover, polypeptides of the present invention can be used to treat disease. For example, patients can be administered a polypeptide of the present invention in an effort to replace absent or decreased levels of the polypeptide (e.g., insulin), to supplement absent or decreased levels of a different polypeptide (e.g., hemoglobin S for hemoglobin B), to inhibit the activity of a polypeptide (e.g., an oncogene), to activate the activity of a polypeptide (e.g., by binding to a receptor), to reduce the activity of a membrane bound receptor by competing with it for free ligand (e.g., soluble TNF receptors used in reducing inflammation), or to bring about a desired response (e.g., blood vessel growth).

Similarly, antibodies directed to a polypeptide of the present invention can also be used to treat disease. For example, administration of an antibody directed to a polypeptide of the present invention can bind and reduce overproduction of the polypeptide. Similarly, administration of an antibody can activate the polypeptide, such as by binding to a polypeptide bound to a membrane (receptor).

At the very least, the polypeptides of the present invention can be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. Polypeptides can also be used to raise antibodies, which in turn are used to measure protein expression from a recombinant cell, as a way of assessing transformation of the host cell. Moreover, the polypeptides of the present invention can be used to test the following biological activities.

Biological Activities

The polynucleotides and polypeptides of the present invention can be used in assays to test for one or more biological activities. If these polynucleotides and polypeptides do exhibit activity in a particular assay, it is likely that these molecules may be involved in the diseases associated with the biological activity. Thus, the polynucleotides and polypeptides could be used to treat the associated disease.

35 Immune Activity

A polypeptide or polynucleotide of the present invention may be useful in treating deficiencies or disorders of the immune system, by activating or inhibiting the

10

15

20

25

30

35

proliferation, differentiation, or mobilization (chemotaxis) of immune cells. Immune cells develop through a process called hematopoiesis, producing myeloid (platelets, red blood cells, neutrophils, and macrophages) and lymphoid (B and T lymphocytes) cells from pluripotent stem cells. The etiology of these immune deficiencies or disorders may be genetic, somatic, such as cancer or some autoimmune disorders, acquired (e.g., by chemotherapy or toxins), or infectious. Moreover, a polynucleotide or polypeptide of the present invention can be used as a marker or detector of a particular immune system disease or disorder.

A polynucleotide or polypeptide of the present invention may be useful in treating or detecting deficiencies or disorders of hematopoietic cells. A polypeptide or polynucleotide of the present invention could be used to increase differentiation and proliferation of hematopoietic cells, including the pluripotent stem cells, in an effort to treat those disorders associated with a decrease in certain (or many) types hematopoietic cells. Examples of immunologic deficiency syndromes include, but are not limited to: blood protein disorders (e.g. agammaglobulinemia, dysgammaglobulinemia), ataxia telangiectasia, common variable immunodeficiency, Digeorge Syndrome, HIV infection, HTLV-BLV infection, leukocyte adhesion deficiency syndrome, lymphopenia, phagocyte bactericidal dysfunction, severe combined immunodeficiency (SCIDs), Wiskott-Aldrich Disorder, anemia, thrombocytopenia, or hemoglobinuria.

Moreover, a polypeptide or polynucleotide of the present invention could also be used to modulate hemostatic (the stopping of bleeding) or thrombolytic activity (clot formation). For example, by increasing hemostatic or thrombolytic activity, a polynucleotide or polypeptide of the present invention could be used to treat blood coagulation disorders (e.g., afibrinogenemia, factor deficiencies), blood platelet disorders (e.g. thrombocytopenia), or wounds resulting from trauma, surgery, or other causes. Alternatively, a polynucleotide or polypeptide of the present invention that can decrease hemostatic or thrombolytic activity could be used to inhibit or dissolve clotting. These molecules could be important in the treatment of heart attacks (infarction), strokes, or scarring.

A polynucleotide or polypeptide of the present invention may also be useful in treating or detecting autoimmune disorders. Many autoimmune disorders result from inappropriate recognition of self as foreign material by immune cells. This inappropriate recognition results in an immune response leading to the destruction of the host tissue. Therefore, the administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing autoimmune disorders.

10

15

20

25

30

Examples of autoimmune disorders that can be treated or detected by the present invention include, but are not limited to: Addison's Disease, hemolytic anemia, antiphospholipid syndrome, rheumatoid arthritis, dermatitis, allergic encephalomyelitis, glomerulonephritis, Goodpasture's Syndrome, Graves' Disease, Multiple Sclerosis, Myasthenia Gravis, Neuritis, Ophthalmia, Bullous Pemphigoid, Pemphigus, Polyendocrinopathies, Purpura, Reiter's Disease, Stiff-Man Syndrome, Autoimmune Thyroiditis, Systemic Lupus Erythematosus, Autoimmune Pulmonary Inflammation, Guillain-Barre Syndrome, insulin dependent diabetes mellitis, and autoimmune inflammatory eye disease.

Similarly, allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems, may also be treated by a polypeptide or polynucleotide of the present invention. Moreover, these molecules can be used to treat anaphylaxis, hypersensitivity to an antigenic molecule, or blood group incompatibility.

A polynucleotide or polypeptide of the present invention may also be used to treat and/or prevent organ rejection or graft-versus-host disease (GVHD). Organ rejection occurs by host immune cell destruction of the transplanted tissue through an immune response. Similarly, an immune response is also involved in GVHD, but, in this case, the foreign transplanted immune cells destroy the host tissues. The administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing organ rejection or GVHD.

Similarly, a polypeptide or polynucleotide of the present invention may also be used to modulate inflammation. For example, the polypeptide or polynucleotide may inhibit the proliferation and differentiation of cells involved in an inflammatory response. These molecules can be used to treat inflammatory conditions, both chronic and acute conditions, including inflammation associated with infection (e.g., septic shock, sepsis, or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine induced lung injury, inflammatory bowel disease, Crohn's disease, or resulting from over production of cytokines (e.g., TNF or IL-1.)

Hyperproliferative Disorders

A polypeptide or polynucleotide can be used to treat or detect hyperproliferative disorders, including neoplasms. A polypeptide or polynucleotide of the present invention may inhibit the proliferation of the disorder through direct or indirect interactions. Alternatively, a polypeptide or polynucleotide of the present invention may proliferate other cells which can inhibit the hyperproliferative disorder.

For example, by increasing an immune response, particularly increasing antigenic qualities of the hyperproliferative disorder or by proliferating, differentiating, or mobilizing T-cells, hyperproliferative disorders can be treated. This immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, decreasing an immune response may also be a method of treating hyperproliferative disorders, such as a chemotherapeutic agent.

Examples of hyperproliferative disorders that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but are not limited to neoplasms located in the: abdomen, bone, breast, digestive system, liver, pancreas, peritoneum, endocrine glands (adrenal, parathyroid, pituitary, testicles, ovary, thymus, thyroid), eye, head and neck, nervous (central and peripheral), lymphatic system, pelvic, skin, soft tissue, spleen, thoracic, and urogenital.

Similarly, other hyperproliferative disorders can also be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of such hyperproliferative disorders include, but are not limited to: hypergammaglobulinemia, lymphoproliferative disorders, paraproteinemias, purpura, sarcoidosis, Sezary Syndrome, Waldenstron's Macroglobulinemia, Gaucher's Disease, histiocytosis, and any other hyperproliferative disease, besides neoplasia, located in an organ system listed above.

Infectious Disease

A polypeptide or polynucleotide of the present invention can be used to treat or detect infectious agents. For example, by increasing the immune response, particularly increasing the proliferation and differentiation of B and/or T cells, infectious diseases may be treated. The immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, the polypeptide or polynucleotide of the present invention may also directly inhibit the infectious agent, without necessarily eliciting an immune response.

Viruses are one example of an infectious agent that can cause disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of viruses, include, but are not limited to the following DNA and RNA viral families: Arbovirus, Adenoviridae, Arenaviridae, Arterivirus, Birnaviridae, Bunyaviridae, Caliciviridae, Circoviridae, Coronaviridae, Flaviviridae, Hepadnaviridae (Hepatitis), Herpesviridae (such as, Cytomegalovirus, Herpes

5

10

15

20

25

30

35

Simplex, Herpes Zoster), Mononegavirus (e.g., Paramyxoviridae, Morbillivirus, Rhabdoviridae), Orthomyxoviridae (e.g., Influenza), Papovaviridae, Parvoviridae, Picornaviridae, Poxviridae (such as Smallpox or Vaccinia), Reoviridae (e.g., Rotavirus), Retroviridae (HTLV-I, HTLV-II, Lentivirus), and Togaviridae (e.g.,

- Rubivirus). Viruses falling within these families can cause a variety of diseases or symptoms, including, but not limited to: arthritis, bronchiollitis, encephalitis, eye infections (e.g., conjunctivitis, keratitis), chronic fatigue syndrome, hepatitis (A, B, C, E, Chronic Active, Delta), meningitis, opportunistic infections (e.g., AIDS), pneumonia, Burkitt's Lymphoma, chickenpox, hemorrhagic fever, Measles, Mumps,
- Parainfluenza, Rabies, the common cold, Polio, leukemia, Rubella, sexually transmitted diseases, skin diseases (e.g., Kaposi's, warts), and viremia. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Similarly, bacterial or fungal agents that can cause disease or symptoms and that

can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following Gram-Negative and Gram-positive bacterial families and fungi: Actinomycetales (e.g., Corynebacterium, Mycobacterium, Norcardia), Aspergillosis, Bacillaceae (e.g., Anthrax, Clostridium), Bacteroidaceae, Blastomycosis, Bordetella, Borrelia, Brucellosis, Candidiasis, Campylobacter,

- Coccidioidomycosis, Cryptococcosis, Dermatocycoses, Enterobacteriaceae (Klebsiella, Salmonella, Serratia, Yersinia), Erysipelothrix, Helicobacter, Legionellosis, Leptospirosis, Listeria, Mycoplasmatales, Neisseriaceae (e.g., Acinetobacter, Gonorrhea, Menigococcal), Pasteurellacea Infections (e.g., Actinobacillus, Heamophilus, Pasteurella), Pseudomonas, Rickettsiaceae, Chlamydiaceae, Syphilis,
- and Staphylococcal. These bacterial or fungal families can cause the following diseases or symptoms, including, but not limited to: bacteremia, endocarditis, eye infections (conjunctivitis, tuberculosis, uveitis), gingivitis, opportunistic infections (e.g., AIDS related infections), paronychia, prosthesis-related infections, Reiter's Disease, respiratory tract infections, such as Whooping Cough or Empyema, sepsis, Lyme
- Disease, Cat-Scratch Disease, Dysentery, Paratyphoid Fever, food poisoning, Typhoid, pneumonia, Gonorrhea, meningitis, Chlamydia, Syphilis, Diphtheria, Leprosy, Paratuberculosis, Tuberculosis, Lupus, Botulism, gangrene, tetanus, impetigo, Rheumatic Fever, Scarlet Fever, sexually transmitted diseases, skin diseases (e.g., cellulitis, dermatocycoses), toxemia, urinary tract infections, wound infections.
- A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Moreover, parasitic agents causing disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following families: Amebiasis, Babesiosis, Coccidiosis, Cryptosporidiosis, Dientamoebiasis, Dourine, Ectoparasitic, Giardiasis, Helminthiasis, Leishmaniasis, Theileriasis, Toxoplasmosis, Trypanosomiasis, and Trichomonas. These parasites can cause a variety of diseases or symptoms, including, but not limited to: Scabies, Trombiculiasis, eye infections, intestinal disease (e.g., dysentery, giardiasis), liver disease, lung disease, opportunistic infections (e.g., AIDS related), Malaria, pregnancy complications, and toxoplasmosis. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Preferably, treatment using a polypeptide or polynucleotide of the present invention could either be by administering an effective amount of a polypeptide to the patient, or by removing cells from the patient, supplying the cells with a polynucleotide of the present invention, and returning the engineered cells to the patient (ex vivo therapy). Moreover, the polypeptide or polynucleotide of the present invention can be used as an antigen in a vaccine to raise an immune response against infectious disease.

Regeneration

15

20

25

30

35

A polynucleotide or polypeptide of the present invention can be used to differentiate, proliferate, and attract cells, leading to the regeneration of tissues. (See, Science 276:59-87 (1997).) The regeneration of tissues could be used to repair, replace, or protect tissue damaged by congenital defects, trauma (wounds, burns, incisions, or ulcers), age, disease (e.g. osteoporosis, osteocarthritis, periodontal disease, liver failure), surgery, including cosmetic plastic surgery, fibrosis, reperfusion injury, or systemic cytokine damage.

Tissues that could be regenerated using the present invention include organs (e.g., pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac), vascular (including vascular endothelium), nervous, hematopoietic, and skeletal (bone, cartilage, tendon, and ligament) tissue. Preferably, regeneration occurs without or decreased scarring. Regeneration also may include angiogenesis.

Moreover, a polynucleotide or polypeptide of the present invention may increase regeneration of tissues difficult to heal. For example, increased tendon/ligament regeneration would quicken recovery time after damage. A polynucleotide or polypeptide of the present invention could also be used prophylactically in an effort to avoid damage. Specific diseases that could be treated include of tendinitis, carpal tunnel syndrome, and other tendon or ligament defects. A further example of tissue

10

20

25

30

BNGDOCID- -WO

regeneration of non-healing wounds includes pressure ulcers, ulcers associated with vascular insufficiency, surgical, and traumatic wounds.

Similarly, nerve and brain tissue could also be regenerated by using a polynucleotide or polypeptide of the present invention to proliferate and differentiate nerve cells. Diseases that could be treated using this method include central and peripheral nervous system diseases, neuropathies, or mechanical and traumatic disorders (e.g., spinal cord disorders, head trauma, cerebrovascular disease, and stoke). Specifically, diseases associated with peripheral nerve injuries, peripheral neuropathy (e.g., resulting from chemotherapy or other medical therapies), localized neuropathies, and central nervous system diseases (e.g., Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome), could all be treated using the polynucleotide or polypeptide of the present invention.

15 Chemotaxis

A polynucleotide or polypeptide of the present invention may have chemotaxis activity. A chemotaxic molecule attracts or mobilizes cells (e.g., monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells) to a particular site in the body, such as inflammation, infection, or site of hyperproliferation. The mobilized cells can then fight off and/or heal the particular trauma or abnormality.

A polynucleotide or polypeptide of the present invention may increase chemotaxic activity of particular cells. These chemotactic molecules can then be used to treat inflammation, infection, hyperproliferative disorders, or any immune system disorder by increasing the number of cells targeted to a particular location in the body. For example, chemotaxic molecules can be used to treat wounds and other trauma to tissues by attracting immune cells to the injured location. Chemotactic molecules of the present invention can also attract fibroblasts, which can be used to treat wounds.

It is also contemplated that a polynucleotide or polypeptide of the present invention may inhibit chemotactic activity. These molecules could also be used to treat disorders. Thus, a polynucleotide or polypeptide of the present invention could be used as an inhibitor of chemotaxis.

Binding Activity

A polypeptide of the present invention may be used to screen for molecules that bind to the polypeptide or for molecules to which the polypeptide binds. The binding of the polypeptide and the molecule may activate (agonist), increase, inhibit

10

15

20

25

30

35

(antagonist), or decrease activity of the polypeptide or the molecule bound. Examples of such molecules include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

Preferably, the molecule is closely related to the natural ligand of the polypeptide, e.g., a fragment of the ligand, or a natural substrate, a ligand, a structural or functional mimetic. (See, Coligan et al., Current Protocols in Immunology 1(2):Chapter 5 (1991).) Similarly, the molecule can be closely related to the natural receptor to which the polypeptide binds, or at least, a fragment of the receptor capable of being bound by the polypeptide (e.g., active site). In either case, the molecule can be rationally designed using known techniques.

Preferably, the screening for these molecules involves producing appropriate cells which express the polypeptide, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or *E. coli*. Cells expressing the polypeptide (or cell membrane containing the expressed polypeptide) are then preferably contacted with a test compound potentially containing the molecule to observe binding, stimulation, or inhibition of activity of either the polypeptide or the molecule.

The assay may simply test binding of a candidate compound to the polypeptide, wherein binding is detected by a label, or in an assay involving competition with a labeled competitor. Further, the assay may test whether the candidate compound results in a signal generated by binding to the polypeptide.

Alternatively, the assay can be carried out using cell-free preparations, polypeptide/molecule affixed to a solid support, chemical libraries, or natural product mixtures. The assay may also simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide, measuring polypeptide/molecule activity or binding, and comparing the polypeptide/molecule activity or binding to a standard.

Preferably, an ELISA assay can measure polypeptide level or activity in a sample (e.g., biological sample) using a monoclonal or polyclonal antibody. The antibody can measure polypeptide level or activity by either binding, directly or indirectly, to the polypeptide or by competing with the polypeptide for a substrate.

All of these above assays can be used as diagnostic or prognostic markers. The molecules discovered using these assays can be used to treat disease or to bring about a particular result in a patient (e.g., blood vessel growth) by activating or inhibiting the polypeptide/molecule. Moreover, the assays can discover agents which may inhibit or enhance the production of the polypeptide from suitably manipulated cells or tissues.

10

15

20

25

35

Therefore, the invention includes a method of identifying compounds which bind to a polypeptide of the invention comprising the steps of: (a) incubating a candidate binding compound with a polypeptide of the invention; and (b) determining if binding has occurred. Moreover, the invention includes a method of identifying agonists/antagonists comprising the steps of: (a) incubating a candidate compound with a polypeptide of the invention, (b) assaying a biological activity, and (b) determining if a biological activity of the polypeptide has been altered.

Other Activities

A polypeptide or polynucleotide of the present invention may also increase or decrease the differentiation or proliferation of embryonic stem cells, besides, as discussed above, hematopoietic lineage.

A polypeptide or polynucleotide of the present invention may also be used to modulate mammalian characteristics, such as body height, weight, hair color, eye color, skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery). Similarly, a polypeptide or polynucleotide of the present invention may be used to modulate mammalian metabolism affecting catabolism, anabolism, processing, utilization, and storage of energy.

A polypeptide or polynucleotide of the present invention may be used to change a mammal's mental state or physical state by influencing biorhythms, caricadic rhythms, depression (including depressive disorders), tendency for violence, tolerance for pain, reproductive capabilities (preferably by Activin or Inhibin-like activity), hormonal or endocrine levels, appetite, libido, memory, stress, or other cognitive qualities.

A polypeptide or polynucleotide of the present invention may also be used as a food additive or preservative, such as to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional components.

30 Other Preferred Embodiments

Other preferred embodiments of the claimed invention include an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 50 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of

10

15

20

25

30

35

positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Clone Sequence and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Start Codon and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Similarly preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 150 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

Further preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 500 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

A further preferred embodiment is a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the nucleotide sequence of SEQ ID NO:X beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence of SEQ ID NO:X.

Also preferred is an isolated nucleic acid molecule which hybridizes under stringent hybridization conditions to a nucleic acid molecule, wherein said nucleic acid molecule which hybridizes does not hybridize under stringent hybridization conditions to a nucleic acid molecule having a nucleotide sequence consisting of only A residues or of only T residues.

Also preferred is a composition of matter comprising a DNA molecule which comprises a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the material deposited with the American Type

10

15

20

25

30

35

Culture Collection and given the ATCC Deposit Number shown in Table 1 for said cDNA Clone Identifier.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in the nucleotide sequence of a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the deposit given the ATCC Deposit Number shown in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said sequence of at least 50 contiguous nucleotides is included in the nucleotide sequence of the complete open reading frame sequence encoded by said human cDNA clone.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 150 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 500 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is a method for detecting in a biological sample a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing a nucleotide sequence of at least one nucleic acid molecule in said sample with a sequence selected from said group and determining whether the sequence of said nucleic acid molecule in said sample is at least 95% identical to said selected sequence.

Also preferred is the above method wherein said step of comparing sequences comprises determining the extent of nucleic acid hybridization between nucleic acid molecules in said sample and a nucleic acid molecule comprising said sequence selected from said group. Similarly, also preferred is the above method wherein said step of comparing sequences is performed by comparing the nucleotide sequence determined from a nucleic acid molecule in said sample with said sequence selected from said group. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

10

15

20

25

30

35

A further preferred embodiment is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting nucleic acid molecules in said sample, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for identifying the species, tissue or cell type of a biological sample can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject nucleic acid molecules, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for diagnosing a pathological condition can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a composition of matter comprising isolated nucleic acid molecules wherein the nucleotide sequences of said nucleic acid molecules comprise a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

10

15

20

25

30

35

DISCOUCID- JAIN DOADTOOM I.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1.

Also preferred is a polypeptide, wherein said sequence of contiguous amino acids is included in the amino acid sequence of SEQ ID NO:Y in the range of positions beginning with the residue at about the position of the First Amino Acid of the Secreted Portion and ending with the residue at about the Last Amino Acid of the Open Reading Frame as set forth for SEQ ID NO:Y in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the complete amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is a polypeptide wherein said sequence of contiguous amino acids is included in the amino acid sequence of a secreted portion of the secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

10

15

20

25

30

35

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is an isolated antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method for detecting in a biological sample a polypeptide comprising an amino acid sequence which is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group and determining whether the sequence of said polypeptide molecule in said sample is at least 90% identical to said sequence of at least 10 contiguous amino acids.

Also preferred is the above method wherein said step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group comprises determining the extent of specific binding of polypeptides in said sample to an antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method wherein said step of comparing sequences is performed by comparing the amino acid sequence determined from a polypeptide molecule in said sample with said sequence selected from said group.

10

15

20

25

30

35

Also preferred is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting polypeptide molecules in said sample, if any, comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO: Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method for identifying the species, tissue or cell type of a biological sample, which method comprises a step of detecting polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the above group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

In any of these methods, the step of detecting said polypeptide molecules includes using an antibody.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a nucleotide sequence encoding a polypeptide wherein said polypeptide comprises an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

10

15

20

25

30

35

Also preferred is an isolated nucleic acid molecule, wherein said nucleotide sequence encoding a polypeptide has been optimized for expression of said polypeptide in a prokaryotic host.

Also preferred is an isolated nucleic acid molecule, wherein said polypeptide comprises an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method of making a recombinant vector comprising

inserting any-of-the above-isolated-nucleic-acid-molecule-into-a-vector. Also preferred is
the recombinant vector produced by this method. Also preferred is a method of making
a recombinant host cell comprising introducing the vector into a host cell, as well as the
recombinant host cell produced by this method.

Also preferred is a method of making an isolated polypeptide comprising culturing this recombinant host cell under conditions such that said polypeptide is expressed and recovering said polypeptide. Also preferred is this method of making an isolated polypeptide, wherein said recombinant host cell is a eukaryotic cell and said polypeptide is a secreted portion of a human secreted protein comprising an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y beginning with the residue at the position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y wherein Y is an integer set forth in Table 1 and said position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y is defined in Table 1; and an amino acid sequence of a secreted portion of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The isolated polypeptide produced by this method is also preferred.

Also preferred is a method of treatment of an individual in need of an increased level of a secreted protein activity, which method comprises administering to such an individual a pharmaceutical composition comprising an amount of an isolated polypeptide, polynucleotide, or antibody of the claimed invention effective to increase the level of said protein activity in said individual.

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

10

Examples '

Example 1: Isolation of a Selected cDNA Clone From the Deposited Sample

Each cDNA clone in a cited ATCC deposit is contained in a plasmid vector. Table 1 identifies the vectors used to construct the cDNA library from which each clone was isolated. In many cases, the vector used to construct the library is a phage vector from which a plasmid has been excised. The table immediately below correlates the related plasmid for each phage vector used in constructing the cDNA library. For example, where a particular clone is identified in Table 1 as being isolated in the vector "Lambda Zap," the corresponding deposited clone is in "pBluescript."

Vector Used to Construct Library		Corresponding Deposited Plasmid		
	Lambda Zap	pBluescript (pBS)		
	Uni-Zap XR	pBluescript (pBS)		
15	Zap Express	pBK		
	lafmid BA	plafmid BA		
	pSport1	pSport1		
	pCMVSport 2.0	pCMVSport 2.0		
	pCMVSport 3.0	pCMVSport 3.0		
20	pCR [®] 2.1	pCR [®] 2.1		
		•		

Vectors Lambda Zap (U.S. Patent Nos. 5,128,256 and 5,286,636), Uni-Zap XR (U.S. Patent Nos. 5,128, 256 and 5,286,636), Zap Express (U.S. Patent Nos. 5,128,256 and 5,286,636), pBluescript (pBS) (Short, J. M. et al., Nucleic Acids Res. 16:7583-7600 (1988); Alting-Mees, M. A. and Short, J. M., Nucleic Acids Res.

- 25 17:9494 (1989)) and pBK (Alting-Mees, M. A. et al., Strategies 5:58-61 (1992)) are commercially available from Stratagene Cloning Systems, Inc., 11011 N. Torrey Pines Road, La Jolla, CA, 92037. pBS contains an ampicillin resistance gene and pBK contains a neomycin resistance gene. Both can be transformed into E. coli strain XL-1 Blue, also available from Stratagene. pBS comes in 4 forms SK+, SK-, KS+ and KS.
- The S and K refers to the orientation of the polylinker to the T7 and T3 primer sequences which flank the polylinker region ("S" is for SacI and "K" is for KpnI which are the first sites on each respective end of the linker). "+" or "-" refer to the orientation of the f1 origin of replication ("ori"), such that in one orientation, single stranded rescue initiated from the f1 ori generates sense strand DNA and in the other, antisense.

Vectors pSport1, pCMVSport 2.0 and pCMVSport 3.0, were obtained from Life Technologies, Inc., P. O. Box 6009, Gaithersburg, MD 20897. All Sport vectors contain an ampicillin resistance gene and may be transformed into E. coli strain

35

15

20

25

30

35

DH10B, also available from Life Technologies. (See, for instance, Gruber, C. E., et al., Focus 15:59 (1993).) Vector lafmid BA (Bento Soares, Columbia University, NY) contains an ampicillin resistance gene and can be transformed into E. coli strain XL-1 Blue. Vector pCR[®]2.1, which is available from Invitrogen, 1600 Faraday Avenue, 5 Carlsbad, CA 92008, contains an ampicillin resistance gene and may be transformed into E. coli strain DH10B, available from Life Technologies. (See, for instance, Clark, J. M., Nuc. Acids Res. 16:9677-9686 (1988) and Mead, D. et al., Bio/Technology 9: (1991).) Preferably, a polynucleotide of the present invention does not comprise the phage vector sequences identified for the particular clone in Table 1, as well as the corresponding plasmid vector sequences designated above.

The deposited material in the sample assigned the ATCC Deposit Number cited in Table 1 for any given cDNA clone also may contain one or more additional plasmids. each comprising a cDNA clone different from that given clone. Thus, deposits sharing the same ATCC Deposit Number contain at least a plasmid for each cDNA clone identified in Table 1. Typically, each ATCC deposit sample cited in Table 1 comprises a mixture of approximately equal amounts (by weight) of about 50 plasmid DNAs, each containing a different cDNA clone; but such a deposit sample may include plasmids for more or less than 50 cDNA clones, up to about 500 cDNA clones.

Two approaches can be used to isolate a particular clone from the deposited sample of plasmid DNAs cited for that clone in Table 1. First, a plasmid is directly isolated by screening the clones using a polynucleotide probe corresponding to SEO ID NO:X.

Particularly, a specific polynucleotide with 30-40 nucleotides is synthesized using an Applied Biosystems DNA synthesizer according to the sequence reported.

The oligonucleotide is labeled, for instance, with ³²P-γ-ATP using T4 polynucleotide kinase and purified according to routine methods. (E.g., Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring, NY (1982).) The plasmid mixture is transformed into a suitable host, as indicated above (such as XL-1 Blue (Stratagene)) using techniques known to those of skill in the art, such as those provided by the vector supplier or in related publications or patents cited above. The transformants are plated on 1.5% agar plates (containing the appropriate selection agent, e.g., ampicillin) to a density of about 150 transformants (colonies) per plate. These plates are screened using Nylon membranes according to routine methods for bacterial colony screening (e.g., Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Edit., (1989), Cold Spring Harbor Laboratory Press, pages 1.93 to 1.104), or other techniques known to those of skill in the art.

10

15

20

25

30

35

Alternatively, two primers of 17-20 nucleotides derived from both ends of the SEQ ID NO:X (i.e., within the region of SEQ ID NO:X bounded by the 5' NT and the 3' NT of the clone defined in Table 1) are synthesized and used to amplify the desired cDNA using the deposited cDNA plasmid as a template. The polymerase chain reaction is carried out under routine conditions, for instance, in 25 µl of reaction mixture with 0.5 ug of the above cDNA template. A convenient reaction mixture is 1.5-5 mM MgCl₂, 0.01% (w/v) gelatin, 20 µM each of dATP, dCTP, dGTP, dTTP, 25 pmol of each primer and 0.25 Unit of Taq polymerase. Thirty five cycles of PCR (denaturation at 94°C for 1 min; annealing at 55°C for 1 min; elongation at 72°C for 1 min) are performed with a Perkin-Elmer Cetus automated thermal cycler. The amplified product is analyzed by agarose gel-electrophoresis and the DNA band with expected molecular weight is excised and purified. The PCR product is verified to be the selected sequence by subcloning and sequencing the DNA product.

Several methods are available for the identification of the 5' or 3' non-coding portions of a gene which may not be present in the deposited clone. These methods include but are not limited to, filter probing, clone enrichment using specific probes, and protocols similar or identical to 5' and 3' "RACE" protocols which are well known in the art. For instance, a method similar to 5' RACE is available for generating the missing 5' end of a desired full-length transcript. (Fromont-Racine et al., Nucleic Acids Res. 21(7):1683-1684 (1993).)

Briefly, a specific RNA oligonucleotide is ligated to the 5' ends of a population of RNA presumably containing full-length gene RNA transcripts. A primer set containing a primer specific to the ligated RNA oligonucleotide and a primer specific to a known sequence of the gene of interest is used to PCR amplify the 5' portion of the desired full-length gene. This amplified product may then be sequenced and used to generate the full length gene.

This above method starts with total RNA isolated from the desired source, although poly-A+ RNA can be used. The RNA preparation can then be treated with phosphatase if necessary to eliminate 5' phosphate groups on degraded or damaged RNA which may interfere with the later RNA ligase step. The phosphatase should then be inactivated and the RNA treated with tobacco acid pyrophosphatase in order to remove the cap structure present at the 5' ends of messenger RNAs. This reaction leaves a 5' phosphate group at the 5' end of the cap cleaved RNA which can then be ligated to an RNA oligonucleotide using T4 RNA ligase.

This modified RNA preparation is used as a template for first strand cDNA synthesis using a gene specific oligonucleotide. The first strand synthesis reaction is

used as a template for PCR amplification of the desired 5' end using a primer specific to the ligated RNA oligonucleotide and a primer specific to the known sequence of the gene of interest. The resultant product is then sequenced and analyzed to confirm that the 5' end sequence belongs to the desired gene.

5

10

15

20

25

30

35

Example 2: Isolation of Genomic Clones Corresponding to a Polynucleotide

A human genomic P1 library (Genomic Systems, Inc.) is screened by PCR using primers selected for the cDNA sequence corresponding to SEQ ID NO:X., according to the method described in Example 1. (See also, Sambrook.)

Example 3: Tissue Distribution of Polypeptide

Tissue distribution of mRNA expression of polynucleotides of the present invention is determined using protocols for Northern blot analysis, described by, among others, Sambrook et al. For example, a cDNA probe produced by the method described in Example 1 is labeled with P³² using the rediprimeTM DNA labeling system (Amersham Life Science), according to manufacturer's instructions. After labeling, the probe is purified using CHROMA SPIN-100TM column (Clontech Laboratories, Inc.), according to manufacturer's protocol number PT1200-1. The purified labeled probe is then used to examine various human tissues for mRNA expression.

Multiple Tissue Northern (MTN) blots containing various human tissues (H) or human immune system tissues (IM) (Clontech) are examined with the labeled probe using ExpressHybTM hybridization solution (Clontech) according to manufacturer's protocol number PT1190-1. Following hybridization and washing, the blots are mounted and exposed to film at -70°C overnight, and the films developed according to standard procedures.

Example 4: Chromosomal Mapping of the Polynucleotides

An oligonucleotide primer set is designed according to the sequence at the 5' end of SEQ ID NO:X. This primer preferably spans about 100 nucleotides. This primer set is then used in a polymerase chain reaction under the following set of conditions: 30 seconds, 95°C; 1 minute, 56°C; 1 minute, 70°C. This cycle is repeated 32 times followed by one 5 minute cycle at 70°C. Human, mouse, and harnster DNA is used as template in addition to a somatic cell hybrid panel containing individual chromosomes or chromosome fragments (Bios, Inc). The reactions is analyzed on

15

20

25

30

35

either 8% polyacrylamide gels or 3.5 % agarose gels. Chromosome mapping is determined by the presence of an approximately 100 bp PCR fragment in the particular somatic cell hybrid.

5 Example 5: Bacterial Expression of a Polypeptide

A polynucleotide encoding a polypeptide of the present invention is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' ends of the DNA sequence, as outlined in Example 1, to synthesize insertion fragments. The primers used to amplify the cDNA insert should preferably contain restriction sites, such as BamHI and XbaI, at the 5' end of the primers in order to clone the amplified product into the expression vector. For example, BamHI and XbaI correspond to the restriction enzyme sites on the bacterial expression vector pQE-9. (Qiagen, Inc., Chatsworth, CA). This plasmid vector encodes antibiotic resistance (Amp^r), a bacterial origin of replication (ori), an IPTG-regulatable promoter/operator (P/O), a ribosome binding site (RBS), a 6-histidine tag (6-His), and restriction enzyme cloning sites.

The pQE-9 vector is digested with BamHI and XbaI and the amplified fragment is ligated into the pQE-9 vector maintaining the reading frame initiated at the bacterial RBS. The ligation mixture is then used to transform the E. coli strain M15/rep4 (Qiagen, Inc.) which contains multiple copies of the plasmid pREP4, which expresses the lacI repressor and also confers kanamycin resistance (Kan^r). Transformants are identified by their ability to grow on LB plates and ampicillin/kanamycin resistant colonies are selected. Plasmid DNA is isolated and confirmed by restriction analysis.

Clones containing the desired constructs are grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 ug/ml) and Kan (25 ug/ml). The O/N culture is used to inoculate a large culture at a ratio of 1:100 to 1:250. The cells are grown to an optical density 600 (O.D.⁶⁰⁰) of between 0.4 and 0.6. IPTG (Isopropyl-B-D-thiogalacto pyranoside) is then added to a final concentration of 1 mM. IPTG induces by inactivating the lacI repressor, clearing the P/O leading to increased gene expression.

Cells are grown for an extra 3 to 4 hours. Cells are then harvested by centrifugation (20 mins at 6000Xg). The cell pellet is solubilized in the chaotropic agent 6 Molar Guanidine HCl by stirring for 3-4 hours at 4°C. The cell debris is removed by centrifugation, and the supernatant containing the polypeptide is loaded onto a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (available from QIAGEN, Inc., *supra*). Proteins with a 6 x His tag bind to the Ni-NTA resin with high

10

15

20

25

30

affinity and can be purified in a simple one-step procedure (for details see: The QIAexpressionist (1995) QIAGEN, Inc., supra).

Briefly, the supernatant is loaded onto the column in 6 M guanidine-HCl, pH 8, the column is first washed with 10 volumes of 6 M guanidine-HCl, pH 8, then washed with 10 volumes of 6 M guanidine-HCl pH 6, and finally the polypeptide is eluted with 6 M guanidine-HCl, pH 5.

The purified protein is then renatured by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the protein can be successfully refolded while immobilized on the Ni-NTA column. The recommended conditions are as follows: renature using a linear 6M-1M urea gradient in 500 mM-NaCl, 20% glycerol, 20 mM Tris/HCl pH 7.4, containing protease inhibitors. The renaturation should be performed over a period of 1.5 hours or more. After renaturation the proteins are eluted by the addition of 250 mM immidazole. Immidazole is removed by a final dialyzing step against PBS or 50 mM sodium acetate pH 6 buffer plus 200 mM NaCl. The purified protein is stored at 4°C or frozen at -80°C.

In addition to the above expression vector, the present invention further includes an expression vector comprising phage operator and promoter elements operatively linked to a polynucleotide of the present invention, called pHE4a. (ATCC Accession Number 209645, deposited on February 25, 1998.) This vector contains: 1) a neomycinphosphotransferase gene as a selection marker, 2) an E. coli origin of replication, 3) a T5 phage promoter sequence, 4) two lac operator sequences, 5) a Shine-Delgarno sequence, and 6) the lactose operon repressor gene (lacIq). The origin of replication (oriC) is derived from pUC19 (LTI, Gaithersburg, MD). The promoter sequence and operator sequences are made synthetically.

DNA can be inserted into the pHEa by restricting the vector with NdeI and XbaI, BamHI, XhoI, or Asp718, running the restricted product on a gel, and isolating the larger fragment (the stuffer fragment should be about 310 base pairs). The DNA insert is generated according to the PCR protocol described in Example 1, using PCR primers having restriction sites for NdeI (5' primer) and XbaI, BamHI, XhoI, or Asp718 (3' primer). The PCR insert is gel purified and restricted with compatible enzymes. The insert and vector are ligated according to standard protocols.

The engineered vector could easily be substituted in the above protocol to express protein in a bacterial system.

10

15

20

25

30

Example 6: Purification of a Polypeptide from an Inclusion Body

The following alternative method can be used to purify a polypeptide expressed in *E coli* when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10°C.

Upon completion of the production phase of the *E. coli* fermentation, the cell culture is cooled to 4-10°C and the cells harvested by continuous centrifugation at 15,000 rpm (Heraeus Sepatech). On the basis of the expected yield of protein per unit weight of cell paste and the amount of purified protein required, an appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM Tris, 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

The cells are then lysed by passing the solution through a microfluidizer (Microfuidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 xg for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000 xg centrifugation for 15 min., the pellet is discarded and the polypeptide containing supernatant is incubated at 4°C overnight to allow further GuHCl extraction.

Following high speed centrifugation (30,000 xg) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA by vigorous stirring. The refolded diluted protein solution is kept at 4°C without mixing for 12 hours prior to further purification steps.

To clarify the refolded polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 µm membrane filter with appropriate surface area (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a stepwise manner. The absorbance at 280 nm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

10

15

20

25

30

35

Fractions containing the polypeptide are then pooled and mixed with 4 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant A₂₈₀ monitoring of the effluent. Fractions containing the polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

The resultant polypeptide should exhibit greater than 95% purity after the above refolding and purification steps. No major contaminant bands should be observed from Commassie blue stained 16% SDS-PAGE gel when 5 µg of purified protein is loaded. The purified protein can also be tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

Example 7: Cloning and Expression of a Polypeptide in a Baculovirus Expression System

In this example, the plasmid shuttle vector pA2 is used to insert a polynucleotide into a baculovirus to express a polypeptide. This expression vector contains the strong polyhedrin promoter of the Autographa californica nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction sites such as BamHI, Xba I and Asp718. The polyadenylation site of the simian virus 40 ("SV40") is used for efficient polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from E. coli under control of a weak Drosophila promoter in the same orientation, followed by the polyadenylation signal of the polyhedrin gene. The inserted genes are flanked on both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate a viable virus that express the cloned polynucleotide.

Many other baculovirus vectors can be used in place of the vector above, such as pAc373, pVL941, and pAcIM1, as one skilled in the art would readily appreciate, as long as the construct provides appropriately located signals for transcription, translation, secretion and the like, including a signal peptide and an in-frame AUG as required. Such vectors are described, for instance, in Luckow et al., Virology 170:31-39 (1989).

10

15

20

25

30

35

Specifically, the cDNA sequence contained in the deposited clone, including the AUG initiation codon and the naturally associated leader sequence identified in Table 1, is amplified using the PCR protocol described in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the pA2 vector does not need a second signal peptide. Alternatively, the vector can be modified (pA2 GP) to include a baculovirus leader sequence, using the standard methods described in Summers et al., "A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures," Texas Agricultural Experimental Station Bulletin No. 1555 (1987).

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The plasmid is digested with the corresponding restriction enzymes and optionally, can be dephosphorylated using calf intestinal phosphatase, using routine procedures known in the art. The DNA is then isolated from a 1% agarose gel using a commercially available kit ("Geneclean" BIO 101 Inc., La Jolla, Ca.).

The fragment and the dephosphorylated plasmid are ligated together with T4 DNA ligase. E. coli HB101 or other suitable E. coli hosts such as XL-1 Blue (Stratagene Cloning Systems, La Jolla, CA) cells are transformed with the ligation mixture and spread on culture plates. Bacteria containing the plasmid are identified by digesting DNA from individual colonies and analyzing the digestion product by gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing.

Five μg of a plasmid containing the polynucleotide is co-transfected with 1.0 μg of a commercially available linearized baculovirus DNA ("BaculoGold™ baculovirus DNA", Pharmingen, San Diego, CA), using the lipofection method described by Felgner et al., Proc. Natl. Acad. Sci. USA 84:7413-7417 (1987). One μg of BaculoGold™ virus DNA and 5 μg of the plasmid are mixed in a sterile well of a microtiter plate containing 50 μl of serum-free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards, 10 μl Lipofectin plus 90 μl Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is then incubated for 5 hours at 27° C. The transfection solution is then removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27° C for four days.

After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, *supra*. An agarose gel with "Blue Gal" (Life

10

15

20

Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10.) After appropriate incubation, blue stained plaques are picked with the tip of a micropipettor (e.g., Eppendorf). The agar containing the recombinant viruses is then resuspended in a microcentrifuge tube containing 200 µl of Grace's medium and the suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4° C.

To verify the expression of the polypeptide, Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus containing the polynucleotide at a multiplicity of infection ("MOI") of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine (available from Life Technologies Inc., Rockville, MD). After 42 hours, 5 μ Ci of 35 S-methionine and 5 μ Ci 35 S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then are harvested by centrifugation. The proteins in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).

Microsequencing of the amino acid sequence of the amino terminus of purified protein may be used to determine the amino terminal sequence of the produced protein.

Example 8: Expression of a Polypeptide in Mammalian Cells

The polypeptide of the present invention can be expressed in a mammalian cell.

A typical mammalian expression vector contains a promoter element, which mediates the initiation of transcription of mRNA, a protein coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription is achieved with the early and late promoters from SV40, the long terminal repeats (LTRs) from Retroviruses, e.g., RSV, HTLVI, HIVI and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter).

Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146), pBC12MI (ATCC 67109), pCMVSport 2.0, and pCMVSport 3.0. Mammalian host cells that could be used

35

10

15

20

25

30

35

include, human Hela, 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, quail QC1-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

Alternatively, the polypeptide can be expressed in stable cell lines containing the polynucleotide integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

The transfected gene can also be amplified to express large amounts of the encoded protein. The DHFR (dihydrofolate reductase) marker is useful in developing cell lines that carry several hundred or even several thousand copies of the gene of interest. (See, e.g., Alt, F. W., et al., J. Biol. Chem. 253:1357-1370 (1978); Hamlin, J. L. and Ma, C., Biochem. et Biophys. Acta, 1097:107-143 (1990); Page, M. J. and Sydenham, M. A., Biotechnology 9:64-68 (1991).) Another useful selection marker is the enzyme glutamine synthase (GS) (Murphy et al., Biochem J. 227:277-279 (1991); Bebbington et al., Bio/Technology 10:169-175 (1992). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the production of proteins.

Derivatives of the plasmid pSV2-dhfr (ATCC Accession No. 37146), the expression vectors pC4 (ATCC Accession No. 209646) and pC6 (ATCC Accession No.209647) contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen et al., Molecular and Cellular Biology, 438-447 (March, 1985)) plus a fragment of the CMV-enhancer (Boshart et al., Cell 41:521-530 (1985).) Multiple cloning sites, e.g., with the restriction enzyme cleavage sites BamHI, XbaI and Asp718, facilitate the cloning of the gene of interest. The vectors also contain the 3' intron, the polyadenylation and termination signal of the rat preproinsulin gene, and the mouse DHFR gene under control of the SV40 early promoter.

Specifically, the plasmid pC6, for example, is digested with appropriate restriction enzymes and then dephosphorylated using calf intestinal phosphates by procedures known in the art. The vector is then isolated from a 1% agarose gel.

A polynucleotide of the present invention is amplified according to the protocol outlined in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the vector does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

10

15

20

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The amplified fragment is then digested with the same restriction enzyme and purified on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. *E. coli* HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC6 using, for instance, restriction enzyme analysis.

Chinese hamster ovary cells lacking an active DHFR gene is used for transfection. Five µg of the expression plasmid pC6 is cotransfected with 0.5 µg of the plasmid pSVneo using lipofectin (Felgner et al., supra). The plasmid pSV2-neo contains a dominant selectable marker, the neo gene from Tn5 encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of metothrexate plus 1 mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1 μ M, 2 μ M, 5 μ M, 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100 - $200 \, \mu M$. Expression of the desired gene product is analyzed, for instance, by SDS-PAGE and Western blot or by reversed phase HPLC analysis.

25

30

35

Example 9: Protein Fusions

The polypeptides of the present invention are preferably fused to other proteins. These fusion proteins can be used for a variety of applications. For example, fusion of the present polypeptides to His-tag, HA-tag, protein A, IgG domains, and maltose binding protein facilitates purification. (See Example 5; see also EP A 394,827; Traunecker, et al., Nature 331:84-86 (1988).) Similarly, fusion to IgG-1, IgG-3, and albumin increases the halflife time in vivo. Nuclear localization signals fused to the polypeptides of the present invention can target the protein to a specific subcellular localization, while covalent heterodimer or homodimers can increase or decrease the activity of a fusion protein. Fusion proteins can also create chimeric molecules having more than one function. Finally, fusion proteins can increase solubility and/or stability of the fused protein compared to the non-fused protein. All of the types of fusion

10

15

proteins described above can be made by modifying the following protocol, which outlines the fusion of a polypeptide to an IgG molecule, or the protocol described in Example 5.

Briefly, the human Fc portion of the IgG molecule can be PCR amplified, using primers that span the 5' and 3' ends of the sequence described below. These primers also should have convenient restriction enzyme sites that will facilitate cloning into an expression vector, preferably a mammalian expression vector.

For example, if pC4 (Accession No. 209646) is used, the human Fc portion can be ligated into the BamHI cloning site. Note that the 3' BamHI site should be destroyed. Next, the vector containing the human Fc portion is re-restricted with BamHI, linearizing the vector, and a polynucleotide of the present invention, isolated by the PCR protocol described in Example 1, is ligated into this BamHI site. Note that the polynucleotide is cloned without a stop codon, otherwise a fusion protein will not be produced.

If the naturally occurring signal sequence is used to produce the secreted protein, pC4 does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

20 Human IgG Fc region:

GGGATCCGGAGCCCAAATCTTCTGACAAA'ACTCACACATGCCCACCGTGCC CAGCACCTGAATTCGAGGGTGCACCGTCAGTCTTCCTCTTCCCCCCAAAACC CAAGGACACCCTCATGATCTCCCGGACTCCTGAGGTCACATGCGTGGTGGT GGACGTAAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACG GCGTGGAGGTGCATAATGCCAAGACAAGCCGCGGGAGGAGCAGTACAAC 25 AGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTG AATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAGCCCTCCCAACCCCC ATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCACAGGT GTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCAGCCT GACCTGCCTGGTCAAAGGCTTCTATCCAAGCGACATCGCCGTGGAGTGGGA 30 GAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCCGTGCTGG ACTCCGACGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCA GGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGC ACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAATGAGTGC 35 GACGGCCGCGACTCTAGAGGAT (SEQ ID NO:1)

10

15

20

25

30

35

Example 10: Production of an Antibody from a Polypeptide

The antibodies of the present invention can be prepared by a variety of methods. (See, Current Protocols, Chapter 2.) For example, cells expressing a polypeptide of the present invention is administered to an animal to induce the production of sera containing polyclonal antibodies. In a preferred method, a preparation of the secreted protein is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

In the most preferred method, the antibodies of the present invention are monoclonal antibodies (or protein binding fragments thereof). Such monoclonal antibodies can be prepared using hybridoma technology. (Köhler et al., Nature 256:495 (1975); Köhler et al., Eur. J. Immunol. 6:511 (1976); Köhler et al., Eur. J. Immunol. 6:292 (1976); Hammerling et al., in: Monoclonal Antibodies and T-Cell Hybridomas, Elsevier, N.Y., pp. 563-681 (1981).) In general, such procedures involve immunizing an animal (preferably a mouse) with polypeptide or, more preferably, with a secreted polypeptide-expressing cell. Such cells may be cultured in any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56°C), and supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100 μg/ml of streptomycin.

The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP2O), available from the ATCC. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands et al. (Gastroenterology 80:225-232 (1981).) The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the polypeptide.

Alternatively, additional antibodies capable of binding to the polypeptide can be produced in a two-step procedure using anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, protein specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to the protein-specific antibody can be blocked by the polypeptide.

25

30

35

Such antibodies comprise anti-idiotypic antibodies to the protein-specific antibody and can be used to immunize an animal to induce formation of further protein-specific antibodies.

It will be appreciated that Fab and F(ab')2 and other fragments of the antibodies of the present invention may be used according to the methods disclosed herein. Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')2 fragments). Alternatively, secreted protein-binding fragments can be produced through the application of recombinant DNA technology or through synthetic chemistry.

10 For in vivo use of antibodies in humans, it may be preferable to use "humanized" chimeric monoclonal antibodies. Such antibodies can be produced using genetic constructs derived from hybridoma cells producing the monoclonal antibodies described above. Methods for producing chimeric antibodies are known in the art. (See, for review, Morrison, Science 229:1202 (1985); Oi et al., BioTechniques 4:214 (1986); Cabilly et al., U.S. Patent No. 4,816,567; Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601533; Robinson et al., WO 8702671; Boulianne et al., Nature 312:643 (1984); Neuberger et al., Nature 314:268 (1985).)

20 Example 11: Production Of Secreted Protein For High-Throughput Screening Assays

The following protocol produces a supernatant containing a polypeptide to be tested. This supernatant can then be used in the Screening Assays described in Examples 13-20.

First, dilute Poly-D-Lysine (644 587 Boehringer-Mannheim) stock solution (1mg/ml in PBS) 1:20 in PBS (w/o calcium or magnesium 17-516F Biowhittaker) for a working solution of 50ug/ml. Add 200 ul of this solution to each well (24 well plates) and incubate at RT for 20 minutes. Be sure to distribute the solution over each well (note: a 12-channel pipetter may be used with tips on every other channel). Aspirate off the Poly-D-Lysine solution and rinse with 1ml PBS (Phosphate Buffered Saline). The PBS should remain in the well until just prior to plating the cells and plates may be poly-lysine coated in advance for up to two weeks.

Plate 293T cells (do not carry cells past P+20) at 2 x 10⁵ cells/well in .5ml DMEM(Dulbecco's Modified Eagle Medium)(with 4.5 G/L glucose and L-glutamine (12-604F Biowhittaker))/10% heat inactivated FBS(14-503F Biowhittaker)/1x Penstrep(17-602E Biowhittaker). Let the cells grow overnight.

The next day, mix together in a sterile solution basin: 300 ul Lipofectamine (18324-012 Gibco/BRL) and 5ml Optimem I (31985070 Gibco/BRL)/96-well plate. With a small volume multi-channel pipetter, aliquot approximately 2ug of an expression vector containing a polynucleotide insert, produced by the methods described in Examples 8 or 9, into an appropriately labeled 96-well round bottom plate. With a multi-channel pipetter, add 50ul of the Lipofectamine/Optimem I mixture to each well. Pipette up and down gently to mix. Incubate at RT 15-45 minutes. After about 20 minutes, use a multi-channel pipetter to add 150ul Optimem I to each well. As a control, one plate of vector DNA lacking an insert should be transfected with each set of transfections.

Preferably, the transfection should be performed by tag-teaming the following tasks. By tag-teaming, hands on time is cut in half, and the cells do not spend too much time on PBS. First, person A aspirates off the media from four 24-well plates of cells, and then person B rinses each well with .5-1ml PBS. Person A then aspirates off PBS rinse, and person B, using a12-channel pipetter with tips on every other channel, adds the 200ul of DNA/Lipofectamine/Optimem I complex to the odd wells first, then to the even wells, to each row on the 24-well plates. Incubate at 37°C for 6 hours.

While cells are incubating, prepare appropriate media, either 1%BSA in DMEM with 1x penstrep, or CHO-5 media (116.6 mg/L of CaCl2 (anhyd); 0.00130 mg/L $\text{CuSO}_4\text{-5H}_2\text{O}$; 0.050 mg/L of $\text{Fe}(\text{NO}_3)_3\text{-9H}_2\text{O}$; 0.417 mg/L of $\text{Fe}\text{SO}_4\text{-7H}_2\text{O}$; 311.80 20 mg/L of Kcl; 28.64 mg/L of MgCl₂; 48.84 mg/L of MgSO₄; 6995.50 mg/L of NaCl; 2400.0 mg/L of NaHCO₃; 62.50 mg/L of NaH₂PO₄-H₂0; 71.02 mg/L of Na₂HPO₄; .4320 mg/L of ZnSO₄-7 H_2 O; .002 mg/L of Arachidonic Acid; 1.022 mg/L of Cholesterol; .070 mg/L of DL-alpha-Tocopherol-Acetate; 0.0520 mg/L of Linoleic 25 Acid; 0.010 mg/L of Linolenic Acid; 0.010 mg/L of Myristic Acid; 0.010 mg/L of Oleic Acid; 0.010 mg/L of Palmitric Acid; 0.010 mg/L of Palmitic Acid; 100 mg/L of Pluronic F-68; 0.010 mg/L of Stearic Acid; 2.20 mg/L of Tween 80; 4551 mg/L of D-Glucose; 130.85 mg/ml of L- Alanine; 147.50 mg/ml of L-Arginine-HCL; 7.50 mg/ml of L-Asparagine-H₂0; 6.65 mg/ml of L-Aspartic Acid; 29.56 mg/ml of L-Cystine-30 2HCL-H₂0; 31.29 mg/ml of L-Cystine-2HCL; 7.35 mg/ml of L-Glutamic Acid; 365.0 mg/ml of L-Glutamine; 18.75 mg/ml of Glycine; 52.48 mg/ml of L-Histidine-HCL-H₂0; 106.97 mg/ml of L-Isoleucine; 111.45 mg/ml of L-Leucine; 163.75 mg/ml of L-Lysine HCL; 32.34 mg/ml of L-Methionine; 68.48 mg/ml of L-Phenylalainine; 40.0 mg/ml of L-Proline; 26.25 mg/ml of L-Serine; 101.05 mg/ml of L-Threonine; 19.22 35 mg/ml of L-Tryptophan; 91.79 mg/ml of L-Tryrosine-2Na-2H,0; 99.65 mg/ml of L-Valine; 0.0035 mg/L of Biotin; 3.24 mg/L of D-Ca Pantothenate; 11.78 mg/L of

20

25

30

35

Choline Chloride; 4.65 mg/L of Folic Acid; 15.60 mg/L of i-Inositol; 3.02 mg/L of Niacinamide; 3.00 mg/L of Pyridoxal HCL; 0.031 mg/L of Pyridoxine HCL; 0.319 mg/L of Riboflavin; 3.17 mg/L of Thiamine HCL; 0.365 mg/L of Thymidine; and 0.680 mg/L of Vitamin B₁₂; 25 mM of HEPES Buffer; 2.39 mg/L of Na Hypoxanthine; 0.105 mg/L of Lipoic Acid; 0.081 mg/L of Sodium Putrescine-2HCL; 55.0 mg/L of Sodium Pyruvate; 0.0067 mg/L of Sodium Selenite; 20uM of Ethanolamine; 0.122 mg/L of Ferric Citrate; 41.70 mg/L of Methyl-B-Cyclodextrin complexed with Linoleic Acid; 33.33 mg/L of Methyl-B-Cyclodextrin complexed with Oleic Acid; and 10 mg/L of Methyl-B-Cyclodextrin complexed with Retinal) with 2mm glutamine and 1x penstrep. (BSA (81-068-3 Bayer) 100gm dissolved in 1L DMEM for a 10% BSA stock solution). Filter the media and collect 50 ul for endotoxin assay in 15ml polystyrene conical.

The transfection reaction is terminated, preferably by tag-teaming, at the end of the incubation period. Person A aspirates off the transfection media, while person B adds 1.5ml appropriate media to each well. Incubate at 37°C for 45 or 72 hours depending on the media used: 1%BSA for 45 hours or CHO-5 for 72 hours.

On day four, using a 300ul multichannel pipetter, aliquot 600ul in one 1ml deep well plate and the remaining supernatant into a 2ml deep well. The supernatants from each well can then be used in the assays described in Examples 13-20.

It is specifically understood that when activity is obtained in any of the assays described below using a supernatant, the activity originates from either the polypeptide directly (e.g., as a secreted protein) or by the polypeptide inducing expression of other proteins, which are then secreted into the supernatant. Thus, the invention further provides a method of identifying the protein in the supernatant characterized by an activity in a particular assay.

Example 12: Construction of GAS Reporter Construct

One signal transduction pathway involved in the differentiation and proliferation of cells is called the Jaks-STATs pathway. Activated proteins in the Jaks-STATs pathway bind to gamma activation site "GAS" elements or interferon-sensitive responsive element ("ISRE"), located in the promoter of many genes. The binding of a protein to these elements alter the expression of the associated gene.

GAS and ISRE elements are recognized by a class of transcription factors called Signal Transducers and Activators of Transcription, or "STATs." There are six members of the STATs family. Stat1 and Stat3 are present in many cell types, as is Stat2 (as response to IFN-alpha is widespread). Stat4 is more restricted and is not in

10

15

20

25

many cell types though it has been found in T helper class I, cells after treatment with IL-12. Stat5 was originally called mammary growth factor, but has been found at higher concentrations in other cells including myeloid cells. It can be activated in tissue culture cells by many cytokines.

The STATs are activated to translocate from the cytoplasm to the nucleus upon tyrosine phosphorylation by a set of kinases known as the Janus Kinase ("Jaks") family. Jaks represent a distinct family of soluble tyrosine kinases and include Tyk2, Jak1, Jak2, and Jak3. These kinases display significant sequence similarity and are generally catalytically inactive in resting cells.

The Jaks are activated by a wide range of receptors summarized in the Table below. (Adapted from review by Schidler and Darnell, Ann. Rev. Biochem. 64:621-51 (1995).) A cytokine receptor family, capable of activating Jaks, is divided into two groups: (a) Class 1 includes receptors for IL-2, IL-3, IL-4, IL-6, IL-7, IL-9, IL-11, IL-12, IL-15, Epo, PRL, GH, G-CSF, GM-CSF, LIF, CNTF, and thrombopoietin; and (b) Class 2 includes IFN-a, IFN-g, and IL-10. The Class 1 receptors share a conserved cysteine motif (a set of four conserved cysteines and one tryptophan) and a WSXWS motif (a membrane proxial region encoding Trp-Ser-Xxx-Trp-Ser (SEQ ID NO:2)).

Thus, on binding of a ligand to a receptor, Jaks are activated, which in turn activate STATs, which then translocate and bind to GAS elements. This entire process is encompassed in the Jaks-STATs signal transduction pathway.

Therefore, activation of the Jaks-STATs pathway, reflected by the binding of the GAS or the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells. For example, growth factors and cytokines are known to activate the Jaks-STATs pathway. (See Table below.) Thus, by using GAS elements linked to reporter molecules, activators of the Jaks-STATs pathway can be identified.

•	Ligand	tyk2	JAKs Jak1	Jak2	Jak3	<u>STATS</u>	GAS(elements) or ISRE
5	IFN family IFN-a/B IFN-g Il-10	+	+ + ?	- + ?	- - 	1,2,3 1 1,3	ISRE GAS (IRF1>Lys6>IFP)
10	gp130 family IL-6 (Pleiotrohic) Il-11(Pleiotrohic) OnM(Pleiotrohic)	· + · · · · · · · · · · · · · · · · · ·	+ + +	+ ? +	? ? ? ?	1,3 1,3 1,3	GAS (IRF1>Lys6>IFP)
15	LIF(Pleiotrohic) CNTF(Pleiotrohic) G-CSF(Pleiotrohic) IL-12(Pleiotrohic)	? -/+ ? +	+ + + + + + + + + + + + + + + + + + + +	+ + ? +	? ? ? +	1,3 1,3 1,3 1,3	
20	g-C family IL-2 (lymphocytes) IL-4 (lymph/myeloid) IL-7 (lymphocytes) IL-9 (lymphocytes) IL-13 (lymphocyte) IL-15	- - - - ?	+ + + + + + + +	- - - ? ?	+ + + + + + +	1,3,5 6 5 5 6 5	GAS GAS (IRF1 = IFP >>Ly6)(IgH) GAS GAS GAS GAS
30	gp140 family IL-3 (myeloid) IL-5 (myeloid) GM-CSF (myeloid)	- - -	- - -	+ + +	<u>-</u> -	5 5 5	GAS (IRF1>IFP>>Ly6) GAS GAS
35	Growth hormone fami GH PRL EPO	? ? ?	- +/- -	+ + +	- -	5 1,3,5 5	GAS(B-CAS>IRF1=IFP>>Ly6)
40	Receptor Tyrosine Kin EGF PDGF CSF-1	<u>nases</u> ? ? ?	+ + +	+ + +	<u>-</u>	1,3 1,3 1,3	GAS (IRF1) GAS (not IRF1)

WO 98/42738 PCT/US98/05311

To construct a synthetic GAS containing promoter element, which is used in the Biological Assays described in Examples 13-14, a PCR based strategy is employed to generate a GAS-SV40 promoter sequence. The 5' primer contains four tandem copies of the GAS binding site found in the IRF1 promoter and previously demonstrated to bind STATs upon induction with a range of cytokines (Rothman et al., Immunity 1:457-468 (1994).), although other GAS or ISRE elements can be used instead. The 5' primer also contains 18bp of sequence complementary to the SV40 early promoter sequence and is flanked with an XhoI site. The sequence of the 5' primer is: 5':GCGCCTCGAGATTTCCCCGAAATCTAGATTTCCCCGAAATGATTTCCCCGAAATGATTTCCCCGAAATGATTTCCCCGAAATGATTTCCCCGAAATGATTTCCCCGAAATGATTTCCCCGAAATGATTTCCCCGAAATGATTTCCCCGAAATGATTTCCCCGAAATGATTTCCCCGAAATGATTTCCCCGAAATGATTTCCCCCGAAATTATCTGCCATCTCAATTAG:3' (SEQ ID NO:3)

The downstream primer is complementary to the SV40 promoter and is flanked with a Hind III site: 5':GCGGCAAGCTTTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)

PCR amplification is performed using the SV40 promoter template present in the B-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI/Hind III and subcloned into BLSK2-. (Stratagene.) Sequencing with forward and reverse primers confirms that the insert contains the following sequence:

5':CTCGAGATTTCCCCGAAATCTAGATTTCCCCGAAATGATTTCCCCGAAATG
ATTTCCCCGAAATATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCC
CTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGC
CCCATGGCTGACTAATTTTTTTTATTTATTTATGCAGAGGCCGAGGCCGCCTCGGC
CTCTGAGCTATTCCAGAAGTAGTGAGGAGGCCTTTTTTGGAGGCCTAGGCTTT
TGCAAAAAGCTT:3' (SEQ ID NO:5)

With this GAS promoter element linked to the SV40 promoter, a GAS:SEAP2 reporter construct is next engineered. Here, the reporter molecule is a secreted alkaline phosphatase, or "SEAP." Clearly, however, any reporter molecule can be instead of SEAP, in this or in any of the other Examples. Well known reporter molecules that can be used instead of SEAP include chloramphenicol acetyltransferase (CAT), luciferase, alkaline phosphatase, B-galactosidase, green fluorescent protein (GFP), or any protein detectable by an antibody.

The above sequence confirmed synthetic GAS-SV40 promoter element is subcloned into the pSEAP-Promoter vector obtained from Clontech using HindIII and XhoI, effectively replacing the SV40 promoter with the amplified GAS:SV40 promoter element, to create the GAS-SEAP vector. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

5

10

15

20

25

30

10

15

20

25

30

35

Thus, in order to generate mammalian stable cell lines expressing the GAS-SEAP reporter, the GAS-SEAP cassette is removed from the GAS-SEAP vector using SalI and NotI, and inserted into a backbone vector containing the neomycin resistance gene, such as pGFP-1 (Clontech), using these restriction sites in the multiple cloning site, to create the GAS-SEAP/Neo vector. Once this vector is transfected into mammalian cells, this vector can then be used as a reporter molecule for GAS binding as described in Examples 13-14.

Other constructs can be made using the above description and replacing GAS with a different promoter sequence. For example, construction of reporter molecules containing NFK-B and EGR promoter sequences are described in Examples 15 and 16. However, many other promoters can be substituted using the protocols described in these Examples. For instance, SRE, IL-2, NFAT, or Osteocalcin promoters can be substituted, alone or in combination (e.g., GAS/NF-KB/EGR, GAS/NF-KB, Il-2/NFAT, or NF-KB/GAS). Similarly, other cell lines can be used to test reporter construct activity, such as HELA (epithelial), HUVEC (endothelial), Reh (B-cell), Saos-2 (osteoblast), HUVAC (aortic), or Cardiomyocyte.

Example 13: High-Throughput Screening Assay for T-cell Activity.

The following protocol is used to assess T-cell activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate T-cells. T-cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The T-cell used in this assay is Jurkat T-cells (ATCC Accession No. TIB-152), although Molt-3 cells (ATCC Accession No. CRL-1552) and Molt-4 cells (ATCC Accession No. CRL-1582) cells can also be used.

Jurkat T-cells are lymphoblastic CD4+ Th1 helper cells. In order to generate stable cell lines, approximately 2 million Jurkat cells are transfected with the GAS-SEAP/neo vector using DMRIE-C (Life Technologies)(transfection procedure described below). The transfected cells are seeded to a density of approximately 20,000 cells per well and transfectants resistant to 1 mg/ml genticin selected. Resistant colonies are expanded and then tested for their response to increasing concentrations of interferon gamma. The dose response of a selected clone is demonstrated.

Specifically, the following protocol will yield sufficient cells for 75 wells containing 200 ul of cells. Thus, it is either scaled up, or performed in multiple to generate sufficient cells for multiple 96 well plates. Jurkat cells are maintained in RPMI + 10% serum with 1%Pen-Strep. Combine 2.5 mls of OPTI-MEM (Life Technologies)

10

15

20

25

30

with 10 ug of plasmid DNA in a T25 flask. Add 2.5 ml OPTI-MEM containing 50 ul of DMRIE-C and incubate at room temperature for 15-45 mins.

During the incubation period, count cell concentration, spin down the required number of cells (10⁷ per transfection), and resuspend in OPTI-MEM to a final concentration of 10⁷ cells/ml. Then add 1ml of 1 x 10⁷ cells in OPTI-MEM to T25 flask and incubate at 37°C for 6 hrs. After the incubation, add 10 ml of RPMI + 15% serum.

The Jurkat:GAS-SEAP stable reporter lines are maintained in RPMI + 10% serum, 1 mg/ml Genticin, and 1% Pen-Strep. These cells are treated with supernatants containing a polypeptide as produced by the protocol described in Example 11.

On the day of treatment with the supernatant, the cells should be washed and resuspended in fresh RPMI + 10% serum to a density of 500,000 cells per ml. The exact number of cells required will depend on the number of supernatants being screened. For one 96 well plate, approximately 10 million cells (for 10 plates, 100 million cells) are required.

Transfer the cells to a triangular reservoir boat, in order to dispense the cells into a 96 well dish, using a 12 channel pipette. Using a 12 channel pipette, transfer 200 ul of cells into each well (therefore adding 100, 000 cells per well).

After all the plates have been seeded, 50 ul of the supernatants are transferred directly from the 96 well plate containing the supernatants into each well using a 12 channel pipette. In addition, a dose of exogenous interferon gamma (0.1, 1.0, 10 ng) is added to wells H9, H10, and H11 to serve as additional positive controls for the assay.

The 96 well dishes containing Jurkat cells treated with supernatants are placed in an incubator for 48 hrs (note: this time is variable between 48-72 hrs). 35 ul samples from each well are then transferred to an opaque 96 well plate using a 12 channel pipette. The opaque plates should be covered (using sellophene covers) and stored at -20°C until SEAP assays are performed according to Example 17. The plates containing the remaining treated cells are placed at 4°C and serve as a source of material for repeating the assay on a specific well if desired.

As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate Jurkat T cells. Over 30 fold induction is typically observed in the positive control wells.

10

15

20

25

30

Example 14: High-Throughput Screening Assay Identifying Myeloid Activity

The following protocol is used to assess myeloid activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate myeloid cells. Myeloid cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The myeloid cell used in this assay is U937, a pre-monocyte cell line, although TF-1, HL60, or KG1 can be used.

To transiently transfect U937 cells with the GAS/SEAP/Neo construct produced in Example 12, a DEAE-Dextran method (Kharbanda et. al., 1994, Cell Growth & Differentiation, 5:259-265) is used. First, harvest 2x10e⁷ U937 cells and wash with PBS. The U937 cells are usually grown in RPMI 1640 medium containing 10% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 mg/ml streptomycin.

Next, suspend the cells in 1 ml of 20 mM Tris-HCl (pH 7.4) buffer containing 0.5 mg/ml DEAE-Dextran, 8 ug GAS-SEAP2 plasmid DNA, 140 mM NaCl, 5 mM KCl, 375 uM Na₂HPO₄.7H₂O, 1 mM MgCl₂, and 675 uM CaCl₂. Incubate at 37°C for 45 min.

Wash the cells with RPMI 1640 medium containing 10% FBS and then resuspend in 10 ml complete medium and incubate at 37°C for 36 hr.

The GAS-SEAP/U937 stable cells are obtained by growing the cells in 400 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 400 ug/ml G418 for couple of passages.

These cells are tested by harvesting $1x10^8$ cells (this is enough for ten 96-well plates assay) and wash with PBS. Suspend the cells in 200 ml above described growth medium, with a final density of $5x10^5$ cells/ml. Plate 200 ul cells per well in the 96-well plate (or $1x10^5$ cells/well).

Add 50 ul of the supernatant prepared by the protocol described in Example 11. Incubate at 37°C for 48 to 72 hr. As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate U937 cells. Over 30 fold induction is typically observed in the positive control wells. SEAP assay the supernatant according to the protocol described in Example 17.

10

15

20

25

30

35

Example 15: High-Throughput Screening Assay Identifying Neuronal Activity.

When cells undergo differentiation and proliferation, a group of genes are activated through many different signal transduction pathways. One of these genes, EGR1 (early growth response gene 1), is induced in various tissues and cell types upon activation. The promoter of EGR1 is responsible for such induction. Using the EGR1 promoter linked to reporter molecules, activation of cells can be assessed.

Particularly, the following protocol is used to assess neuronal activity in PC12 cell lines. PC12 cells (rat phenochromocytoma cells) are known to proliferate and/or differentiate by activation with a number of mitogens, such as TPA (tetradecanoyl phorbol acetate), NGF (nerve growth factor), and EGF (epidermal growth factor). The EGR1 gene expression is activated during this treatment. Thus, by stably transfecting PC12 cells with a construct containing an EGR promoter linked to SEAP reporter, activation of PC12 cells can be assessed.

The EGR/SEAP reporter construct can be assembled by the following protocol. The EGR-1 promoter sequence (-633 to +1)(Sakamoto K et al., Oncogene 6:867-871 (1991)) can be PCR amplified from human genomic DNA using the following primers:

- 5' GCGCTCGAGGGATGACAGCGATAGAACCCCGG -3' (SEQ ID NO:6)
- 5' GCGAAGCTTCGCGACTCCCGGATCCGCCTC-3' (SEQ ID NO:7)

Using the GAS:SEAP/Neo vector produced in Example 12, EGR1 amplified product can then be inserted into this vector. Linearize the GAS:SEAP/Neo vector using restriction enzymes XhoI/HindIII, removing the GAS/SV40 stuffer. Restrict the EGR1 amplified product with these same enzymes. Ligate the vector and the EGR1 promoter.

To prepare 96 well-plates for cell culture, two mls of a coating solution (1:30 dilution of collagen type I (Upstate Biotech Inc. Cat#08-115) in 30% ethanol (filter sterilized)) is added per one 10 cm plate or 50 ml per well of the 96-well plate, and allowed to air dry for 2 hr.

PC12 cells are routinely grown in RPMI-1640 medium (Bio Whittaker) containing 10% horse serum (JRH BIOSCIENCES, Cat. # 12449-78P), 5% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 ug/ml streptomycin on a precoated 10 cm tissue culture dish. One to four split is done every three to four days. Cells are removed from the plates by scraping and resuspended with pipetting up and down for more than 15 times.

Transfect the EGR/SEAP/Neo construct into PC12 using the Lipofectamine protocol described in Example 11. EGR-SEAP/PC12 stable cells are obtained by growing the cells in 300 ug/ml G418. The G418-free medium is used for routine

10

15

20

25

30

growth but every one to two months, the cells should be re-grown in 300 ug/ml G418 for couple of passages.

To assay for neuronal activity, a 10 cm plate with cells around 70 to 80% confluent is screened by removing the old medium. Wash the cells once with PBS (Phosphate buffered saline). Then starve the cells in low serum medium (RPMI-1640 containing 1% horse serum and 0.5% FBS with antibiotics) overnight.

The next morning, remove the medium and wash the cells with PBS. Scrape off the cells from the plate, suspend the cells well in 2 ml low serum medium. Count the cell number and add more low serum medium to reach final cell density as $5x10^5$ cells/ml.

Add 200 ul of the cell suspension to each well of 96-well plate (equivalent to $1x10^5$ cells/well). Add 50 ul supernatant produced by Example 11, 37°C for 48 to 72 hr. As a positive control, a growth factor known to activate PC12 cells through EGR can be used, such as 50 ng/ul of Neuronal Growth Factor (NGF). Over fifty-fold induction of SEAP is typically seen in the positive control wells. SEAP assay the supernatant according to Example 17.

Example 16: High-Throughput Screening Assay for T-cell Activity

NF-κB (Nuclear Factor κB) is a transcription factor activated by a wide variety of agents including the inflammatory cytokines IL-1 and TNF, CD30 and CD40, lymphotoxin-alpha and lymphotoxin-beta, by exposure to LPS or thrombin, and by expression of certain viral gene products. As a transcription factor, NF-κB regulates the expression of genes involved in immune cell activation, control of apoptosis (NF-κB appears to shield cells from apoptosis), B and T-cell development, anti-viral and antimicrobial responses, and multiple stress responses.

In non-stimulated conditions, NF- κB is retained in the cytoplasm with I-κB (Inhibitor κB). However, upon stimulation, I- κB is phosphorylated and degraded, causing NF- κB to shuttle to the nucleus, thereby activating transcription of target genes. Target genes activated by NF- κB include IL-2, IL-6, GM-CSF, ICAM-1 and class 1 MHC.

Due to its central role and ability to respond to a range of stimuli, reporter constructs utilizing the NF-kB promoter element are used to screen the supernatants produced in Example 11. Activators or inhibitors of NF-kB would be useful in treating

10

diseases. For example, inhibitors of NF-kB could be used to treat those diseases related to the acute or chronic activation of NF-kB, such as rheumatoid arthritis.

The downstream primer is complementary to the 3' end of the SV40 promoter and is flanked with a Hind III site:

5':GCGGCAAGCTTTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)

PCR amplification is performed using the SV40 promoter template present in the pB-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI and Hind III and subcloned into BLSK2-. (Stratagene)

Sequencing with the T7 and T3 primers confirms the insert contains the following

Sequencing with the T7 and T3 primers confirms the insert contains the following sequence:

5':CTCGAGGGACTTTCCCGGGGACTTTCCGGGACTTTCC
ATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCA

20 TCCCGCCCCTAACTCCGCCCAGTTCCGCCCCATGCTGACT
AATTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTC
CAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTT:
3' (SEQ ID NO:10)

Next, replace the SV40 minimal promoter element present in the pSEAP2promoter plasmid (Clontech) with this NF-κB/SV40 fragment using XhoI and HindIII.
However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

In order to generate stable mammalian cell lines, the NF-κB/SV40/SEAP

cassette is removed from the above NF-κB/SEAP vector using restriction enzymes Sall and NotI, and inserted into a vector containing neomycin resistance. Particularly, the NF-κB/SV40/SEAP cassette was inserted into pGFP-1 (Clontech), replacing the GFP gene, after restricting pGFP-1 with SalI and NotI.

10

15

20

25

Once NF-kB/SV40/SEAP/Neo vector is created, stable Jurkat T-cells are created and maintained according to the protocol described in Example 13. Similarly, the method for assaying supernatants with these stable Jurkat T-cells is also described in Example 13. As a positive control, exogenous TNF alpha (0.1,1, 10 ng) is added to wells H9, H10, and H11, with a 5-10 fold activation typically observed.

Example 17: Assay for SEAP Activity

As a reporter molecule for the assays described in Examples 13-16, SEAP activity is assayed using the Tropix Phospho-light Kit (Cat. BP-400) according to the following general procedure. The Tropix Phospho-light Kit supplies the Dilution, Assay, and Reaction Buffers used below.

Prime a dispenser with the 2.5x Dilution Buffer and dispense 15 μ l of 2.5x dilution buffer into Optiplates containing 35 μ l of a supernatant. Seal the plates with a plastic sealer and incubate at 65°C for 30 min. Separate the Optiplates to avoid uneven heating.

Cool the samples to room temperature for 15 minutes. Empty the dispenser and prime with the Assay Buffer. Add 50 µl Assay Buffer and incubate at room temperature 5 min. Empty the dispenser and prime with the Reaction Buffer (see the table below). Add 50 µl Reaction Buffer and incubate at room temperature for 20 minutes. Since the intensity of the chemiluminescent signal is time dependent, and it takes about 10 minutes to read 5 plates on luminometer, one should treat 5 plates at each time and start the second set 10 minutes later.

Read the relative light unit in the luminometer. Set H12 as blank, and print the results. An increase in chemiluminescence indicates reporter activity.

Reaction Buffer Formulation:

Meaction D	unci l'ormulation.	
# of plates	Rxn buffer diluent (ml)	CSPD (ml)
10	60	3
11	65	3.25
12	70	3.5
13	75	3.75
14	80	4
15	85	4.25
16	90	4.5
17	95	4.75
18	100	5
19	105	5.25
20	110	5.5
21	115	5.75
22	120	6

23 125 6.25 24 130 6.5 25 135 6.75 26 140 7 27 145 7.25 28 150 7.5 29 155 7.75 30 160 8 31 165 8.25 32 170 8.5 33 175 8.75 34 180 9 35 185 9.25 36 190 9.5 37 195 9.75 38 200 10 39 205 10.25 40 210 10.5 41 215 10.75 42 220 11 43 225 11.25 44 230 11.5 45 235 11.75 46 240 12 47 245 12.25 48 250 12.5 49 255 12.75 <		•	•
25 135 6.75 26 140 7 27 145 7.25 28 150 7.5 29 155 7.75 30 160 8 31 165 8.25 32 170 8.5 33 175 8.75 34 180 9 35 185 9.25 36 190 9.5 37 195 9.75 38 200 10 39 205 10.25 40 210 10.5 41 215 10.75 42 220 11 43 225 11.25 44 230 11.5 45 235 11.75 46 240 12 47 245 12.25 48 250 12.5 49 255 12.75	23	125	6.25
26 140 7 27 145 7.25 28 150 7.5 29 155 7.75 30 160 8 31 165 8.25 32 170 8.5 33 175 8.75 34 180 9 35 185 9.25 36 190 9.5 37 195 9.75 38 200 10 39 205 10.25 40 210 10.5 41 215 10.75 42 220 11 43 225 11.25 44 230 11.5 45 235 11.75 46 240 12 47 245 12.25 48 250 12.5 49 255 12.75	24	130	6.5
27 145 7.25 28 150 7.5 29 155 7.75 30 160 8 31 165 8.25 32 170 8.5 33 175 8.75 34 180 9 35 185 9.25 36 190 9.5 37 195 9.75 38 200 10 39 205 10.25 40 210 10.5 41 215 10.75 42 220 11 43 225 11.25 44 230 11.5 45 235 11.75 46 240 12 47 245 12.25 48 250 12.5 49 255 12.75	25	135	6.75
28 150 7.5 29 155 7.75 30 160 8 31 165 8.25 32 170 8.5 33 175 8.75 34 180 9 35 185 9.25 36 190 9.5 37 195 9.75 38 200 10 39 205 10.25 40 210 10.5 41 215 10.75 42 220 11 43 225 11.25 44 230 11.5 45 235 11.75 46 240 12 47 245 12.25 48 250 12.5 49 255 12.75	26	140	7
29 155 7.75 30 160 8 31 165 8.25 32 170 8.5 33 175 8.75 34 180 9 35 185 9.25 36 190 9.5 37 195 9.75 38 200 10 39 205 10.25 40 210 10.5 41 215 10.75 42 220 11 43 225 11.25 44 230 11.5 45 235 11.75 46 240 12 47 245 12.25 48 250 12.5 49 255 12.75	27	145	7.25
29 155 7.75 30 160 8 31 165 8.25 32 170 8.5 33 175 8.75 34 180 9 35 185 9.25 36 190 9.5 37 195 9.75 38 200 10 39 205 10.25 40 210 10.5 41 215 10.75 42 220 11 43 225 11.25 44 230 11.5 45 235 11.75 46 240 12 47 245 12.25 48 250 12.5 49 255 12.75	28	150	7.5
31 165 8.25 32 170 8.5 33 175 8.75 34 180 9 35 185 9.25 36 190 9.5 37 195 9.75 38 200 10 39 205 10.25 40 210 10.5 41 215 10.75 42 220 11 43 225 11.25 44 230 11.5 45 235 11.75 46 240 12 47 245 12.25 48 250 12.5 49 255 12.75	29	155	7.75
32 170 8.5 33 175 8.75 34 180 9 35 185 9.25' 36 190 9.5 37 195 9.75 38 200 10 39 205 10.25 40 210 10.5 41 215 10.75 42 220 11 43 225 11.25 44 230 11.5 45 235 11.75 46 240 12 47 245 12.25 48 250 12.5 49 255 12.75	30	160	' 8
33 175 8.75 34 180 9 35 185 9.25 36 190 9.5 37 195 9.75 38 200 10 39 205 10.25 40 210 10.5 41 215 10.75 42 220 11 43 225 11.25 44 230 11.5 45 235 11.75 46 240 12 47 245 12.25 48 250 12.5 49 255 12.75	31	165	8.25
34 180 9 35 185 9.25¹ 36 190 9.5 37 195 9.75 38 200 10 39 205 10.25 40 210 10.5 41 215 10.75 42 220 11 43 225 11.25 44 230 11.5 45 235 11.75 46 240 12 47 245 12.25 48 250 12.5 49 255 12.75	32	170	8.5
34 180 9 35 185 9.25 36 190 9.5 37 195 9.75 38 200 10 39 205 10.25 40 210 10.5 41 215 10.75 42 220 11 43 225 11.25 44 230 11.5 45 235 11.75 46 240 12 47 245 12.25 48 250 12.5 49 255 12.75	33	175	
36 190 9.5 37 195 9.75 38 200 10 39 205 10.25 40 210 10.5 41 215 10.75 42 220 11 43 225 11.25 44 230 11.5 45 235 11.75 46 240 12 47 245 12.25 48 250 12.5 49 255 12.75	34	180	9
37 195 9.75 38 200 10 39 205 10.25 40 210 10.5 41 215 10.75 42 220 11 43 225 11.25 44 230 11.5 45 235 11.75 46 240 12 47 245 12.25 48 250 12.5 49 255 12.75	35	185	9.25
38 200 10 39 205 10.25 40 210 10.5 41 215 10.75 42 220 11 43 225 11.25 44 230 11.5 45 235 11.75 46 240 12 47 245 12.25 48 250 12.5 49 255 12.75	36	190	9.5
39 205 10.25 40 210 10.5 41 215 10.75 42 220 11 43 225 11.25 44 230 11.5 45 235 11.75 46 240 12 47 245 12.25 48 250 12.5 49 255 12.75	37		9.75
40 210 10.5 41 215 10.75 42 220 11 43 225 11.25 44 230 11.5 45 235 11.75 46 240 12 47 245 12.25 48 250 12.5 49 255 12.75	38		10
41 215 10.75 42 220 11 43 225 11.25 44 230 11.5 45 235 11.75 46 240 12 47 245 12.25 48 250 12.5 49 255 12.75			
42 220 11 43 225 11.25 44 230 11.5 45 235 11.75 46 240 12 47 245 12.25 48 250 12.5 49 255 12.75	40		10.5
43 225 11.25 44 230 11.5 45 235 11.75 46 240 12 47 245 12.25 48 250 12.5 49 255 12.75			
44 230 11.5 45 235 11.75 46 240 12 47 245 12.25 48 250 12.5 49 255 12.75			
45 235 46 240 47 245 48 250 49 255 11.75 12.25 12.5 12.75			
46 240 12 47 245 12.25 48 250 12.5 49 255 12.75			
47 245 12.25 48 250 12.5 49 255 12.75			
48 250 12.5 49 255 12.75			
49 255 12.75			
50 260 13			
	50	260	13

Example 18: High-Throughput Screening Assay Identifying Changes in Small Molecule Concentration and Membrane Permeability

Binding of a ligand to a receptor is known to alter intracellular levels of small molecules, such as calcium, potassium, sodium, and pH, as well as alter membrane potential. These alterations can be measured in an assay to identify supernatants which bind to receptors of a particular cell. Although the following protocol describes an assay for calcium, this protocol can easily be modified to detect changes in potassium, sodium, pH, membrane potential, or any other small molecule which is detectable by a fluorescent probe.

The following assay uses Fluorometric Imaging Plate Reader ("FLIPR") to measure changes in fluorescent molecules (Molecular Probes) that bind small molecules. Clearly, any fluorescent molecule detecting a small molecule can be used instead of the calcium fluorescent molecule, fluo-3, used here.

For adherent cells, seed the cells at 10,000 -20,000 cells/well in a Co-star black 96-well plate with clear bottom. The plate is incubated in a CO₂ incubator for 20 hours. The adherent cells are washed two times in Biotek washer with 200 ul of HBSS (Hank's Balanced Salt Solution) leaving 100 ul of buffer after the final wash.

5

10

A stock solution of 1 mg/ml fluo-3 is made in 10% pluronic acid DMSO. To load the cells with fluo-3, 50 ul of 12 ug/ml fluo-3 is added to each well. The plate is incubated at 37°C in a CO₂ incubator for 60 min. The plate is washed four times in the Biotek washer with HBSS leaving 100 ul of buffer.

For non-adherent cells, the cells are spun down from culture media. Cells are re-suspended to 2-5x10⁶ cells/ml with HBSS in a 50-ml conical tube. 4 ul of 1 mg/ml fluo-3 solution in 10% pluronic acid DMSO is added to each ml of cell suspension. The tube is then placed in a 37°C water bath for 30-60 min. The cells are washed twice with HBSS, resuspended to 1x10⁶ cells/ml, and dispensed into a microplate, 100 ul/well. The plate is centrifuged at 1000 rpm for 5 min. The plate is then washed once in Denley CellWash with 200 ul, followed by an aspiration step to 100 ul final volume.

For a non-cell based assay, each well contains a fluorescent molecule, such as fluo-3. The supernatant is added to the well, and a change in fluorescence is detected.

To measure the fluorescence of intracellular calcium, the FLIPR is set for the following parameters: (1) System gain is 300-800 mW; (2) Exposure time is 0.4 second; (3) Camera F/stop is F/2; (4) Excitation is 488 nm; (5) Emission is 530 nm; and (6) Sample addition is 50 ul. Increased emission at 530 nm indicates an extracellular signaling event which has resulted in an increase in the intracellular Ca++ concentration.

20

5

10

15

Example 19: High-Throughput Screening Assay Identifying Tyrosine Kinase Activity

The Protein Tyrosine Kinases (PTK) represent a diverse group of transmembrane and cytoplasmic kinases. Within the Receptor Protein Tyrosine Kinase RPTK) group are receptors for a range of mitogenic and metabolic growth factors including the PDGF, FGF, EGF, NGF, HGF and Insulin receptor subfamilies. In addition there are a large family of RPTKs for which the corresponding ligand is unknown. Ligands for RPTKs include mainly secreted small proteins, but also membrane-bound and extracellular matrix proteins.

30

35

25

Activation of RPTK by ligands involves ligand-mediated receptor dimerization, resulting in transphosphorylation of the receptor subunits and activation of the cytoplasmic tyrosine kinases. The cytoplasmic tyrosine kinases include receptor associated tyrosine kinases of the src-family (e.g., src, yes, lck, lyn, fyn) and non-receptor linked and cytosolic protein tyrosine kinases, such as the Jak family, members of which mediate signal transduction triggered by the cytokine superfamily of receptors (e.g., the Interleukins, Interferons, GM-CSF, and Leptin).

Because of the wide range of known factors capable of stimulating tyrosine kinase activity, the identification of novel human secreted proteins capable of activating tyrosine kinase signal transduction pathways are of interest. Therefore, the following protocol is designed to identify those novel human secreted proteins capable of activating the tyrosine kinase signal transduction pathways.

Seed target cells (e.g., primary keratinocytes) at a density of approximately 25,000 cells per well in a 96 well Loprodyne Silent Screen Plates purchased from Nalge Nunc (Naperville, IL). The plates are sterilized with two 30 minute rinses with 100% ethanol, rinsed with water and dried overnight. Some plates are coated for 2 hr with 100 ml of cell culture grade type I collagen (50 mg/ml), gelatin (2%) or polylysine (50 mg/ml), all of which can be purchased from Sigma Chemicals (St. Louis, MO) or 10% Matrigel purchased from Becton Dickinson (Bedford,MA), or calf serum, rinsed with PBS and stored at 4°C. Cell growth on these plates is assayed by seeding 5,000 cells/well in growth medium and indirect quantitation of cell number through use of alamarBlue as described by the manufacturer Alamar Biosciences, Inc. (Sacramento, CA) after 48 hr. Falcon plate covers #3071 from Becton Dickinson (Bedford,MA) are used to cover the Loprodyne Silent Screen Plates. Falcon Microtest III cell culture plates can also be used in some proliferation experiments.

To prepare extracts, A431 cells are seeded onto the nylon membranes of Loprodyne plates (20,000/200ml/well) and cultured overnight in complete medium. Cells are quiesced by incubation in serum-free basal medium for 24 hr. After 5-20 minutes treatment with EGF (60ng/ml) or 50 ul of the supernatant produced in Example 11, the medium was removed and 100 ml of extraction buffer ((20 mM HEPES pH 7.5, 0.15 M NaCl, 1% Triton X-100, 0.1% SDS, 2 mM Na3VO4, 2 mM Na4P2O7 and a cocktail of protease inhibitors (# 1836170) obtained from Boeheringer Mannheim (Indianapolis, IN) is added to each well and the plate is shaken on a rotating shaker for 5 minutes at 4°C. The plate is then placed in a vacuum transfer manifold and the extract filtered through the 0.45 mm membrane bottoms of each well using house vacuum. Extracts are collected in a 96-well catch/assay plate in the bottom of the vacuum manifold and immediately placed on ice. To obtain extracts clarified by centrifugation, the content of each well, after detergent solubilization for 5 minutes, is removed and centrifuged for 15 minutes at 4°C at 16,000 x g.

Test the filtered extracts for levels of tyrosine kinase activity. Although many methods of detecting tyrosine kinase activity are known, one method is described here.

Generally, the tyrosine kinase activity of a supernatant is evaluated by determining its ability to phosphorylate a tyrosine residue on a specific substrate (a

5

10

15

20

25

30

10

25

biotinylated peptide). Biotinylated peptides that can be used for this purpose include PSK1 (corresponding to amino acids 6-20 of the cell division kinase cdc2-p34) and PSK2 (corresponding to amino acids 1-17 of gastrin). Both peptides are substrates for a range of tyrosine kinases and are available from Boehringer Mannheim.

The tyrosine kinase reaction is set up by adding the following components in order. First, add 10ul of 5uM Biotinylated Peptide, then 10ul ATP/Mg₂₊ (5mM ATP/50mM MgCl₂), then 10ul of 5x Assay Buffer (40mM imidazole hydrochloride, pH7.3, 40 mM beta-glycerophosphate, 1mM EGTA, 100mM MgCl₂, 5 mM MnCl₂, 0.5 mg/ml BSA), then 5ul of Sodium Vanadate(1mM), and then 5ul of water. Mix the components gently and preincubate the reaction mix at 30°C for 2 min. Initial the reaction by adding 10ul of the control enzyme or the filtered supernatant.

The tyrosine kinase assay reaction is then terminated by adding 10 ul of 120mm EDTA and place the reactions on ice.

Tyrosine kinase activity is determined by transferring 50 ul aliquot of reaction

mixture to a microtiter plate (MTP) module and incubating at 37°C for 20 min. This allows the streptavadin coated 96 well plate to associate with the biotinylated peptide. Wash the MTP module with 300ul/well of PBS four times. Next add 75 ul of anti-phospotyrosine antibody conjugated to horse radish peroxidase(anti-P-Tyr-POD(0.5u/ml)) to each well and incubate at 37°C for one hour. Wash the well as above.

Next add 100ul of peroxidase substrate solution (Boehringer Mannheim) and incubate at room temperature for at least 5 mins (up to 30 min). Measure the absorbance of the sample at 405 nm by using ELISA reader. The level of bound peroxidase activity is quantitated using an ELISA reader and reflects the level of tyrosine kinase activity.

Example 20: High-Throughput Screening Assay Identifying Phosphorylation Activity

As a potential alternative and/or compliment to the assay of protein tyrosine kinase activity described in Example 19, an assay which detects activation (phosphorylation) of major intracellular signal transduction intermediates can also be used. For example, as described below one particular assay can detect tyrosine phosphorylation of the Erk-1 and Erk-2 kinases. However, phosphorylation of other molecules, such as Raf, JNK, p38 MAP, Map kinase kinase (MEK), MEK kinase, Src, Muscle specific kinase (MuSK), IRAK, Tec, and Janus, as well as any other

phosphoserine, phosphotyrosine, or phosphothreonine molecule, can be detected by substituting these molecules for Erk-1 or Erk-2 in the following assay.

Specifically, assay plates are made by coating the wells of a 96-well ELISA plate with 0.1ml of protein G (1ug/ml) for 2 hr at room temp, (RT). The plates are then rinsed with PBS and blocked with 3% BSA/PBS for 1 hr at RT. The protein G plates are then treated with 2 commercial monoclonal antibodies (100ng/well) against Erk-1 and Erk-2 (1 hr at RT) (Santa Cruz Biotechnology). (To detect other molecules, this step can easily be modified by substituting a monoclonal antibody detecting any of the above described molecules.) After 3-5 rinses with PBS, the plates are stored at 4°C until use.

A431 cells are seeded at 20,000/well in a 96-well Loprodyne filterplate and cultured overnight in growth medium. The cells are then starved for 48 hr in basal medium (DMEM) and then treated with EGF (6ng/well) or 50 ul of the supernatants obtained in Example 11 for 5-20 minutes. The cells are then solubilized and extracts filtered directly into the assay plate.

After incubation with the extract for 1 hr at RT, the wells are again rinsed. As a positive control, a commercial preparation of MAP kinase (10ng/well) is used in place of A431 extract. Plates are then treated with a commercial polyclonal (rabbit) antibody (1ug/ml) which specifically recognizes the phosphorylated epitope of the Erk-1 and Erk-2 kinases (1 hr at RT). This antibody is biotinylated by standard procedures. The bound polyclonal antibody is then quantitated by successive incubations with Europium-streptavidin and Europium fluorescence enhancing reagent in the Wallac DELFIA instrument (time-resolved fluorescence). An increased fluorescent signal over background indicates a phosphorylation.

25

30

5

10

15

20

Example 21: Method of Determining Alterations in a Gene Corresponding to a Polynucleotide

RNA isolated from entire families or individual patients presenting with a phenotype of interest (such as a disease) is be isolated. cDNA is then generated from these RNA samples using protocols known in the art. (See, Sambrook.) The cDNA is then used as a template for PCR, employing primers surrounding regions of interest in SEQ ID NO:X. Suggested PCR conditions consist of 35 cycles at 95°C for 30 seconds; 60-120 seconds at 52-58°C; and 60-120 seconds at 70°C, using buffer solutions described in Sidransky, D., et al., Science 252:706 (1991).

PCR products are then sequenced using primers labeled at their 5' end with T4 polynucleotide kinase, employing SequiTherm Polymerase. (Epicentre Technologies).

10

15

20

25

30

35

The intron-exon borders of selected exons is also determined and genomic PCR products analyzed to confirm the results. PCR products harboring suspected mutations is then cloned and sequenced to validate the results of the direct sequencing.

PCR products is cloned into T-tailed vectors as described in Holton, T.A. and Graham, M.W., Nucleic Acids Research, 19:1156 (1991) and sequenced with T7 polymerase (United States Biochemical). Affected individuals are identified by mutations not present in unaffected individuals.

Genomic rearrangements are also observed as a method of determining alterations in a gene corresponding to a polynucleotide. Genomic clones isolated according to Example 2 are nick-translated with digoxigenindeoxy-uridine 5'-triphosphate (Boehringer Manheim), and FISH performed as described in Johnson, Cg. et al., Methods Cell Biol. 35:73-99 (1991). Hybridization with the labeled probe is carried out using a vast excess of human cot-1 DNA for specific hybridization to the corresponding genomic locus.

Chromosomes are counterstained with 4,6-diamino-2-phenylidole and propidium iodide, producing a combination of C- and R-bands. Aligned images for precise mapping are obtained using a triple-band filter set (Chroma Technology, Brattleboro, VT) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) and variable excitation wavelength filters. (Johnson, Cv. et al., Genet. Anal. Tech. Appl., 8:75 (1991).) Image collection, analysis and chromosomal fractional length measurements are performed using the ISee Graphical Program System. (Inovision Corporation, Durham, NC.) Chromosome alterations of the genomic region hybridized by the probe are identified as insertions, deletions, and translocations. These alterations are used as a diagnostic marker for an associated disease.

Example 22: Method of Detecting Abnormal Levels of a Polypeptide in a Biological Sample

A polypeptide of the present invention can be detected in a biological sample, and if an increased or decreased level of the polypeptide is detected, this polypeptide is a marker for a particular phenotype. Methods of detection are numerous, and thus, it is understood that one skilled in the art can modify the following assay to fit their particular needs.

For example, antibody-sandwich ELISAs are used to detect polypeptides in a sample, preferably a biological sample. Wells of a microtiter plate are coated with specific antibodies, at a final concentration of 0.2 to 10 ug/ml. The antibodies are either monoclonal or polyclonal and are produced by the method described in Example 10.

10

15

20

25

30

35

The wells are blocked so that non-specific binding of the polypeptide to the well is reduced.

The coated wells are then incubated for > 2 hours at RT with a sample containing the polypeptide. Preferably, serial dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled water to remove unbounded polypeptide.

Next, 50 ul of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature. The plates are again washed three times with deionized or distilled water to remove unbounded conjugate.

Add 75 ul of 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (NPP) substrate solution to each well and incubate 1 hour at room temperature. Measure the reaction by a microtiter plate reader. Prepare a standard curve, using serial dilutions of a control sample, and plot polypeptide concentration on the X-axis (log scale) and fluorescence or absorbance of the Y-axis (linear scale). Interpolate the concentration of the polypeptide in the sample using the standard curve.

Example 23: Formulating a Polypeptide

The secreted polypeptide composition will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the secreted polypeptide alone), the site of delivery, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

As a general proposition, the total pharmaceutically effective amount of secreted polypeptide administered parenterally per dose will be in the range of about 1 μ g/kg/day to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the secreted polypeptide is typically administered at a dose rate of about 1 μ g/kg/hour to about 50 μ g/kg/hour, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

Pharmaceutical compositions containing the secreted protein of the invention are administered orally, rectally, parenterally, intracistemally, intravaginally,

10

15

20

25

30

35

intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), bucally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

The secreted polypeptide is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semi-permeable polymer matrices in the form of shaped articles, e.g., films, or mirocapsules. Sustained-release matrices include polylactides (U.S. Pat. No. 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman, U. et al., Biopolymers 22:547-556 (1983)), poly (2- hydroxyethyl methacrylate) (R. Langer et al., J. Biomed. Mater. Res. 15:167-277 (1981), and R. Langer, Chem. Tech. 12:98-105 (1982)), ethylene vinyl acetate (R. Langer et al.) or poly-D- (-)-3-hydroxybutyric acid (EP 133,988). Sustained-release compositions also include liposomally entrapped polypeptides. Liposomes containing the secreted polypeptide are prepared by methods known per se: DE 3,218,121; Epstein et al., Proc. Natl. Acad. Sci. USA 82:3688-3692 (1985); Hwang et al., Proc. Natl. Acad. Sci. USA 77:4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324. Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal secreted polypeptide therapy.

For parenteral administration, in one embodiment, the secreted polypeptide is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to polypeptides.

Generally, the formulations are prepared by contacting the polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, manose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

The secreted polypeptide is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts.

Any polypeptide to be used for therapeutic administration can be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

Polypeptides ordinarily will be stored in unit or multi-dose containers, for example, sealed ampoules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1% (w/v) aqueous polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized polypeptide using bacteriostatic Water-for-Injection.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the polypeptides of the present invention may be employed in conjunction with other therapeutic compounds.

30

5

10

15

20

10

15

20

25

30

RNSDOCID: WO 0842729A1 I

Example 24: Method of Treating Decreased Levels of the Polypeptide

It will be appreciated that conditions caused by a decrease in the standard or normal expression level of a secreted protein in an individual can be treated by administering the polypeptide of the present invention, preferably in the secreted form. Thus, the invention also provides a method of treatment of an individual in need of an increased level of the polypeptide comprising administering to such an individual a pharmaceutical composition comprising an amount of the polypeptide to increase the activity level of the polypeptide in such an individual.

For example, a patient with decreased levels of a polypeptide receives a daily dose 0.1-100 ug/kg of the polypeptide for six consecutive days. Preferably, the polypeptide is in the secreted form. The exact details of the dosing scheme, based on administration and formulation, are provided in Example 23.

Example 25: Method of Treating Increased Levels of the Polypeptide

Antisense technology is used to inhibit production of a polypeptide of the present invention. This technology is one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer.

For example, a patient diagnosed with abnormally increased levels of a polypeptide is administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5, 2.0 and 3.0 mg/kg day for 21 days. This treatment is repeated after a 7-day rest period if the treatment was well tolerated. The formulation of the antisense polynucleotide is provided in Example 23.

Example 26: Method of Treatment Using Gene Therapy

One method of gene therapy transplants fibroblasts, which are capable of expressing a polypeptide, onto a patient. Generally, fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of the flask and fresh media (e.g., Ham's F12 media, with 10% FBS, penicillin and streptomycin) is added. The flasks are then incubated at 37°C for approximately one week.

WO 98/42738 PCT/US98/05311

161

At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks.

pMV-7 (Kirschmeier, P.T. et al., DNA, 7:219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

The cDNA encoding a polypeptide of the present invention can be amplified using PCR primers which correspond to the 5' and 3' end sequences respectively as set forth in Example 1. Preferably, the 5' primer contains an EcoRI site and the 3' primer includes a HindIII-site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The ligation mixture is then used to transform bacteria HB101, which are then plated onto agar containing kanamycin for the purpose of confirming that the vector has the gene of interest properly inserted.

The amphotropic pA317 or GP+am12 packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the gene is then added to the media and the packaging cells transduced with the vector. The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells).

Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media, containing the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the media from the producer cells. This media is removed and replaced with fresh media. If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is very low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether protein is produced.

The engineered fibroblasts are then transplanted onto the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads.

30

5

10

15

20

It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples. Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

The entire disclosure of each document cited (including patents, patent applications, journal articles, abstracts, laboratory manuals, books, or other disclosures) in the Background of the Invention, Detailed Description, and Examples is hereby incorporated herein by reference.

```
(1) GENERAL INFORMATION:
           (i) APPLICANT: Human Genome Sciences, Inc. et al.
           (ii) TITLE OF INVENTION: 87 Human Secreted Proteins
5
           (iii) NUMBER OF SEQUENCES: 323
           (iv) CORRESPONDENCE ADDRESS:
                 (A) ADDRESSEE: Human Genome Sciences, Inc.
                 (B) STREET: 9410 Key West Avenue
10
                 (C) CITY: Rockville
                 (D) STATE: Maryland
                 (E) COUNTRY: USA
                 (F) ZIP: 20850
15
           (v) COMPUTER READABLE FORM:
                 (A) MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
                 (B) COMPUTER: HP Vectra 486/33
20
                 (C) OPERATING SYSTEM: MSDOS version 6.2
                 (D) SOFTWARE: ASCII Text
           (vi) CURRENT APPLICATION DATA:
25
                 (A) APPLICATION NUMBER:
                 (B) FILING DATE: March 19, 1998
                 (C) CLASSIFICATION:
30
           (vii) PRIOR APPLICATION DATA:
                 (A) APPLICATION NUMBER:
                 (B) FILING DATE:
35
           (viii) ATTORNEY/AGENT INFORMATION:
                 (A) NAME: A. Anders Brookes
                 (B) REGISTRATION NUMBER: 36,373
                 (C) REFERENCE/DOCKET NUMBER: PZ004PCT
40
           (vi) TELECOMMUNICATION INFORMATION:
                 (A) TELEPHONE: (301) 309-8504
                 (B) TELEFAX: (301) 309-8439
45
     (2) INFORMATION FOR SEQ ID NO: 1:
           (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 733 base pairs
50
                  (B) TYPE: nucleic acid
                  (C) STRANDEDNESS: double
                  (D) TOPOLOGY: linear
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
55
```

	GGGATCCGGA GCCCAAATCT TCTGACAAAA CTCACACATG CCCACCGTGC CCAGCACCTG	60
	AATTOGAGGG TGCACCGTCA GTCTTCCTCT TCCCCCCAAA ACCCAAGGAC ACCCTCATGA	120
5	TCTCCCGGAC TCCTGAGGTC ACATGCGTGG TGGTGGACGT AAGCCACGAA GACCCTGAGG	180
	TCAAGITCAA CTGGTACGTG GACGGCGTGG AGGTGCATAA TGCCAAGACA AAGCCGCGGG	240
10	AGGAGCAGTA CAACAGCACG TACCGTGTGG TCAGCGTCCTT CACCGTCCTG CACCAGGACT	300
	GGCTGAATGG CAAGGAGTAC AAGTGCAAGG TCTCCAACAA AGCCCTCCCA ACCCCCATCG	360
	AGAAAACCAT CTCCAAAGCC AAAGGGCAGC CCCGAGAACC ACAGGTGTAÇ ACCCTGCCCC	420
15	CATCCCGGGA TGAGCTGACC AAGAACCAGG TCAGCCTGAC CTGCCTGGTC AAAGGCTTCT	480
	ATCCAAGCGA CATCGCCGTG GAGTGGGAGA GCAATGGGCA GCCGGAGAAC AACTACAAGA	540
20	CCACGCCTCC CGTGCTGGAC TCCGACGGCT CCTTCTTCCT CTACAGCAAG CTCACCGTGG	600
	ACAAGAGCAG GTGGCAGCAG GGGAACGTCT TCTCATGCTC CGTGATGCAT GAGGCTCTGC	660
	ACAACCACTA CACGCAGAAG AGCCTCTCCC TGTCTCCGGG TAAATGAGTG CGACGGCCGC	720
25	GACTCTAGAG GAT	733
30	(2) INFORMATION FOR SEQ ID NO: 2:	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 5 amino acids	
35	(B) TYPE: amino acid (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:	
40	Trp Ser Xaa Trp Ser	
40	1 5	
45	(2) INFORMATION FOR SEQ ID NO: 3:	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 86 base pairs	
50	(B) TYPE: nucleic acid	
<i>-</i>	(C) STRANDEDNESS: double (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:	
55	GCGCCTCGAG ATTTCCCCGA AATCTAGATT TCCCCGAAAT GATTTCCCCG AAATGATTTC	60
	CCCGAAATAT CTGCCATCTC AATTAG	86

	(2) INFORMATION FOR SEQ ID NO: 4:	
5	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 27 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
10	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:	
		22
	GCGGCAAGCT TTTTGCAAAG CCTAGGC	27
15		
	(2) INFORMATION FOR SEQ ID NO: 5:	
20	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 271 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
25	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:	
	CTCGAGATTT CCCCGAAATC TAGATTTCCC CGAAATGATT TCCCCGAAAT GATTTCCCCG	60
20	AAATATCTGC CATCTCAATT AGTCAGCAAC CATAGTCCCG CCCCTAACTC CGCCCATCCC	120
30	GCCCCTAACT CCGCCCAGTT CCGCCCCATTC TCCGCCCCAT GGCTGACTAA TTTTTTTTAT	180
	TTATGCAGAG GCCGAGGCCG CCTCGGCCTC TGAGCTATTC CAGAAGTAGT GAGGAGGCTT	240
35	TTTTGGAGGC CTAGGCTTTT GCAAAAAGCT T	271
40	(2) INFORMATION FOR SEQ ID NO: 6: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 32 base pairs	
45	(B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:	
50	GCGCTCGAGG GATGACAGCG ATAGAACCCC GG	32
55	(2) INFORMATION FOR SEQ ID NO: 7:	
	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 31 base pairs	
60	(B) TYPE: nucleic acid	
~ 1	(c) containerness, double	

240

(D)	TOPOLOGY:	linear
-----	-----------	--------

	(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:	
5	GCGAAGCTTC GCGACTCCCCC GGATCCCCCT C	31
10	(2) INFORMATION FOR SEQ ID NO: 8:	
15	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 12 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
20	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:	
	GGGGACTTTC CC	12
25	(2) INFORMATION FOR SEQ ID NO: 9/:	
30	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 73 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
35	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:	
	GCGGCCTCGA GGGGACTTTC CCGGGGACTT TCCGGGGACT TTCCATCCTG CCATCTCAAT TAG	60
10	CCATCTCAAT TAG	73
	(2) INFORMATION FOR SEQ ID NO: 10:	
15	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 256 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
50	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:	
	CTCGAGGGGA CTTTCCCGGG GACTTTCCGG GGACTTTCCA TCTGCCATCT	60
55	CAATTAGTCA GCAACCATAG TCCCGCCCCT AACTCCGCCC ATCCCGCCCC TAACTCCGCC	120
	CAGTICCGCC CATTCTCCGC CCCATGGCTG ACTAATTTTT TITATTTATG CAGAGGCCGA	180

GGCCGCCTCG GCCTCTGAGC TATTCCAGAA GTAGTGAGGA GGCTTTTTTG GAGGCCTAGG

PCT/US98/05311 WO 98/42738

167

256 CTTTTGCAAA AAGCTT

5

10

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1679 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

15 20 25

30

35

40

45

50

GCAGCGCACC CGGGCGATCG CTTCACGGAT GCGGACGACG TAGCCATCCT TACCTACGTG 60 AAGGAAAATG CCCGCTCGCC CAGCTCCGTC ACCGGTAACG CCTTGTGGAA AGCGATGGAG 120 AAGAGCTCGC TCACGCAGCA CTCGTGGCAG TCCCTGAAGG ACCGCTACCT CAAGCACCTG 180 COGGGCCAGG AGCATAAGTA CCTGCTGGGG GACGCGCCGG TGAGCCCCCTC CTCCCAGAAG 240 CTCAAGCGGA AGGCGGAGGA GGACCCGGAG GCCGCGGATA GCGGGGAACC ACAGAATAAG 300 AGAACTCCAG ATTTGCCTGA AGAAGAGTAT GTGAAGGAAG AAATCCAGGA GAATGAAGAA 360 CCAGTCAAAA AGATGCTTGT GGAAGCCACC CGGGAGTTTG AGGAGGTTGT GGTGGATGAG 420 AGCCCTCCTG ATTTTGAAAT ACATATAACT ATGTGTGATG ATGATCCACC CACACCTGAG 480 540 GAAGACTCAG AAACACAGCC TGATGAGGAG GAAGAAGAAG AAGAAGAAAA AGTTTCTCAA CCAGAGGTGG GAGCTGCCAT TAAGATCATT CGGCAGTTAA TGGAGAAGTT TAACTTGGAT CTATCAACAG TTACACAGGC CTTCCTAAAA AATAGTGGTG AGCTGGAGGC TACTTCCGCC TTCTTAGCGT CTGGTCAGAG AGCTGATGGA TATCCCATTT GGTCCCGACA AGATGACATA 720 GATTTGCAAA AAGATGATGA GGATACCAGA GAGGCATTGG TCAAAAAATT TGGTGCTCAG 780 840 AAAGTCATGG TAGGTGAGGT GGTTAAAAAA AATTGTGACC AATGAACTTT AGAGAGTTCT 900 TGCATTGGAA CTGGCACTTA TTTTCTGACC ATCGCTGCTG TTGCTCTGTG AGTCCTAGAT 960 TTTTGTAGCC AAGCAGAGTT GTAGAGGGGG ATAAAAAGAA AAGAAATTGG ATGTATTTAC 1020 1080 AGCTGTCCTT GAACAAGTAT CAATGTGTTT ATGAAAGGAA GATCTAAATC AGACAGGAGT TGGTCTACAT AGTAGTAATC CATTGTTGGA ATGGAACCCT TGCTATAGTA GTGACAAAGT 1140

55

GAAAGGAAAT TTAGGAGGCA TAGGCCATTT CAGGCAGCAT AAGTAATCTC CTGTCCTTTG GCAGAAGCTC CTTTAGATTG GGATAGATTC CAAATAAAGA ATCTAGAAAT AGGAGAAGAT 1260 TTAATTATGA GGCCTTGAAC ACGGATTATC CCCAAACCCT TGTCATTTCC CCCAGTGAGC 1320

1200

TCTGATTTCT AGACTGCTTT GAAAATGCTG TATTCATTTT GCTAACTTAG TATTTGGGTA 1380 60

	CCCTGCTCTT	TGGCTGTTCT	TTTTTTGGAG	CCCTTCTCAG	TCAAGTCTGC	CGGATGTCTT	1440
5	TCTTTACCTA	CCCCTCAGTT	TTCCTTAAAA	CGCGCACACA	ACTCTAGAGA	GTGTTAAGAA	1500
	TAATGTTACT	TGGTTAATGT	GTTATTTATT	GAGTATTGTT	TGTGCTAAGC	ATTGTGTTAG	1560
	ATTTAAAAAA	TTAGTGGATT	GACTCCACTT	TCTTCTCTTC	TTTTCATTGT	TGAAAATAAA	1620
10	TATAACTTTG	TATTCGAAAA	ааааааааа	AAAATNRCTG	CGGNCCGACA	AGGGAATTC	1679

15 (2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1830 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double

 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

		· -		. 552 10	. 12.	*	
25	GCGACCGCGC	CCTTCAGCTA	GCTCGCTCGC	TCGCTCTGCT	TCCCTGCTGC	CGGCTGCGCA	60
	TGGCTTNGGC	GIIGCÇGCCG	CTGGCGGCGG	TCGAGCNGCC	TGCGSAGCCG	GTACCAGCAG	120
30	TTGCAGAATG	AAGAAGAGTC	TGGAGAACCT	GAACAGGCTG	CAGGTGATGC	TCCTCCACCT	180
	TACAGCAGCA	TTTCTGCAGA	GAGCGCACAT	NATITIGACT	ACAAGGATGA	GICTGGGITT	240
	CCAAAGCCCC	CATCTTACAA	TGTAGCTACA	ACACTGCCCA	GTTATGATGA	AGCGGAGAGG	300
35	ACCAAGGCTG	AAGCTACTAT	CCCTTTGGTT	CCTGGGAGAG	ATGAGGATTT	TGTGGGTCGG	360
	GATGATTTTG	ATGATGCTGA	CCAGCTGAGG	ATAGGAAATG	ATGGGATTTT	CATGITAACT	420
40	TTTTTCATGG	CATTCCTCTT	TAACTGGATT	GGGTTTTTCC	TGTCTTTTTG	CCTGACCACT	480
	TCAGCTGCAG	GAAGGTATGG	GCCCATTTCA	GGATTTGGTC	TCTCTCTAAT	TAAATGGATC	540
	CTGATTGTCA	GGTTTTCCAC	CTATTTCCCT	GGATATTTTG	ATGGTCAGTA	CTCCCTCTCC	600
45	TGGGTGTTCC	TTGTTTTAGG	CTTTCTCCTG	TTTCTCAGAG	GATTTATCAA	TTATGCAAAA	660
	GTTCGGAAGA	TGCCAGAAAC	TTTCTCAAAT	CTCCCCAGGA	CCAGAGTTCT	CTTTATTTAT	720
50	TAAAGATGTT	TTCTGGCAAA	GCCCTTCCTG	CATTTATGAA	TTCTCTCTCA	AGAAGCAAGA	780
	GAACACCTGC	AGGAAGTGAA	TCAAGATGCA	GAACACAGAG	GAATAATCAC	CTGCTTTAAA	840
	AAAATAAAGT	ACTGTTGAAA	AGATCATTTC	TCTCTATTTG	TTCCTAGGTG	TAAAATTTTA	900
55	ATAGTTAATG	CAGAATTCTG	TAATCATTGA	ATCATTAGTG	GTTAATGTTT	GAAAAAGCTC	960
	TTGCAATCAA	GTCTGTGATG	TATTAATAAT	GCCTTATATA	TTGTTTGTAG	TCATTTTAAG	1020
60	TAGCATGAGC	CATGTCCCTG	TAGTCGGTAG	GGGGCAGTCT	TGCTTTATTC	ATCCTCCATC	1080

	TCAAAATGAA	CTTGGAATTA	aatattgtaa	GATATGTATA	ATGCTGGCCA	TTTTAAAGGG	1140
	GTTTTCTCAA	AAGTTAAACT	TTTGTTATGA	CIGIGITITI	GCACATAATC	CATATTTGCT	1200
5	GTTCAAGTTA	ATCTAGAAAT	TTATTCAATT	CIGTATGAAC	ACCTGGAAGC	AAAATCATAG	1260
	TGCAAAAATA	CATTTAAGGT	GTGGTCAAAA	ATAAGTCTTT	AATTGGTAAA	TAATAAGCAT	1320
10	TAATTTTTTA	TAGCCTGTAT	TCACAATTCT	GCGCTACCTT	ATTGTACCTA	AGGGATTCTA	1380
10	AAGGTGTTGT	CACTGTATAA	AACAGAAAGC	ACTAGGATAC	AAATGAAGCT	TAATTACTAA	1440
	AATGTAATTC	TTGACACTCT	TTCTATAATT	AGCGTTCTTC	ACCCCACCC	CCACCCCCAC	1500
15	CCCCCTTATT	TICCTITIGT	CTCCTGGTGA	TTAGGCCAAA	ÇTCTGGGAGT	AAGGAGAGGA	1560
	TTAGGTACTT	AGGAGCAAAG	AAAGAAGTAG	CTTGGAACTT	TTGAGATGAT	CCCTAACATA	1620
20	CTGTACTACT	TGCTTTTACA	ATGTGTTAGC	AGAAACCAGT	GGGTTATAAT	GTAGAATGAT	1680
20	GTGCTTTCTG	CCCAAGTGGT	AATTCATCTT	GGTTTGCTAT	GTTAAAACTG	TAAATACAAC	1740
	AGAACATTAA	TAAATATCTC	TTGTGTAGCA	CCTTTTAAAA	AAAAAAAA	AAAÁAAAAA	1800
25	AAAAAAAAA	AANCCCGGGG	GGGGGCCCCIN		f		1830

30 (2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1212 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

40	TGTTTGAAGT	TGTTACTTTT	GTTTACAGCA	AAGTITGATG	TAGTGTGCAG	TAGTGAGCTC	60
	TAGACTGATC	ТТТТТСТААА	TCAGAAAGTG	attaaagtat	GCACAACCAA	AGGCAGGTTT	120
45	TTCTTTTTCA	TTTATTCAGC	AACTATTTAT	TAAGCATCAA	CTCTGTGCCA	GGCACGTTAC	180
43	TAGCTGCTAC	ATACTGTCTG	AACATGACAT	ACGGTTAAGT	AACTTTACAA	TTATTATCAA	240
	ATACTTCAAT	GTAGATATTT	CTTAAGTTGA	AATAGCATTA	ACTAGGATAA	TGCTTTCATG	300
50	TTATTTTATT	TGTCTTGTGA	TAGAAATTCA	ACTITICTACC	ATCTTAAAAC	TAGGITGCTA	360
	TAAAAATAGG	AGGATGAAGT	CAATAAAGTT	TATGCCAGIT	TAAAAACTGG	AAGGAAAAGG	420
55	TAAGAGCTCT	CCATTATAAA	ATAGITGCAT	TCGGTTAATT	TTTACACATT	AGTGCATTGC	480
22	GTATATCAAC	TGGCCCTCAA	TGAAGCATTT	AAGTGCTTGG	AATTITACTA	AACTGACTTT	540
	TTTGCAACTT	TGGGAGATTT	TTGAGGGGAG	TGTTGAAAAT	TGCCAAACAC	TCACCTCTTA	600
60	СПСАВАВСТТ	САВАТАВАЯТ	ACACATTITC	AAGAGGGAGC	ACCITTTATA	TTTGATAAGT	660

	TTTCATTATA	AACCTTATAA	TACCAGTCAC	AAAGAGGTTG	TCTGTCTATG	GTTTAGCAAA	720
5	CATTICCITT	TCTTTTTGGA	AGTGTGATTG	CAATTGCAGA	ACAGAAAGTG	AGAAAACACT	780
	GCCAGCGGTG	ATTGCTACTT	GAGGTAGTTT	TTTACAACTA	CCATTTCCCC	TCCATGAAAT	840
	TATGTGAAAT	TTATTTTATC	TTTGGGAAAA	GTTGAGAÁGA	TAGTAAAAGA	ATTAGGAATT	900
10	TAAAATTACA	GGGAAAAATA	TOTAAGTGAA	AAGCAATAAA	TATTTTGTTC	ACTITGCTAT	960
	CAAGATGTTC	ACTATCAGAT	ATTTATTATA	TGGCAGCAAT	TTATATTTT	AATCATTGCC	1020
15	CATTAATAGA	CGCAGTAAAA	TATTTTTGAA	TCAGACATTT	GCCCTTCTA	TGTGCATTAA	1080
	AATTGTCTTT	TGTACTGTAA	GTTACTGTTA	ATTTGAATAT	TTTATTGAAC	TETETECETG	1140
	TGCCTTTATA	ATATAAAGTT	GTTTCTACAA	CTTTTAATGA	TCTTAATAAA	GAATACTTTA	1200
20	AGAAAAAAA	AA					1212
						1	

25 (2) INFORMATION FOR SEQ ID NO: 14:

30

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2061 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14: .

35	GCTTTTCCTC	CGACTTCCGG	ACATCTCCCT	GGGAGTCGCG	CAGAGTGGAG	TCAAAGGCAA	60
	CCAGTGCTCG	CTGCGGTCTC	TGGGGATCGG	GACCGCGCG	cccccccc	AGCGGGATGT	120
40	TCCGGGGCTT	GAGCAGTTGG	TTGGGCTTGC	AGCAGCCGGT	GGCAGGCGGT	GGGCAGCCCA	180
	ATGGAGATGC	TCCACCCGAG	CAGCCGTCCG	AGACGGTGGC	TGAGTCTGCG	GAGGAGGAGC	240
	TGCAGCAAGC	GGGAGACCAG	GAGCTCCTCC	ACCAGGCCAA	AGACTTCGGC	AACTATTTAT	300
45	TTAACTTTGC	ATCTGCTGCC	ACAAAAAAGA	TAACTGAATC	AGTTGCTGAA	ACAGCACAAA	360
	CAATAAAGAA	ATCCGTAGAA	GAAGGAAAAA	TAGATGGCAT	CATTGACAAG	ACAATTATAG	420
50	GAGATTTTCA	GAAGGAACAG	DTTTAAAAAA	TTGAAGAGCA	ACATACAAAG	AAGTCAGAAG	480
	CAGCTGTGCC	CCCATGGGTT	GACACTAACG	ATGAAGAAAC	AATTCAACAA	CAAATTTTGG	540
	CCTTATCAGC	TGACAAGAGG	AATTTCCTTC	GTGACCCTCC	GCTGCCTG	CAATTTAATT	600
55	TCGACTTTGA	TCAGATGTAC	CCCGTGGCCC	TGGTCATGCT	CCAGGAGGAT	GAGCTGCTAR	660
	CAAGATGAGA	TTTGCCCTCG	TTCCTAAACT	TGTGAAGGAA	GAAGTGTTCT	GGAGGAACTA	720
60	CTTTTACCGC	GTCTCCCTGA	TTAAGCAGTC	AGCCCAGCTC	ACGGCCCTCG	CTGCCCAACA	780

	GCAGGCCGCA	GGGAAGGGAG	GAGAAGAGCA	ATGGCÅGAGA	GCAAGATTTG	CCGCTGGAGA	840
	GGCAGTACGG	CCCAAAACGC	CACCCGTTGT	AATCAAATCT	CAGCTTAAAA	CTCAAGAGGA	900
5	TGAGGAAGAA	ATTTCTACTA	GCCCAGGTGT	TTCTGAGTTT	GTCAGTGATG	CCTTCGATGC	960
	CTGTAACCTA	AATCAGGAAG	ATCTAAGGAA	AGAAATGGAG	CAACTAGTGC	TTGACAAAAA	1020
	GCAAGAGGAG	ACAGCCGTAC	TGGAAGAGGA	TTCTGCAGAT	TGGGAAAAAG	AACTGCAGCA	1080
10	GGAACTICAA	GAATATGAAG	TGGTGACAGA	ATCTGAAAAA	CGAGATGAAA	ACTGGGATAA	1140
	GGAAATAGAG	AAAATGCTTC	AAGAGGAAAA	TTAGCTGTTC	CTGAAATAGA	AGAATAATCC	1200
15	TTAACAGTCT	GCAAACTGAC	ATTAAATTCT	AGATGTTGAC	AATTACTGAA	TCAGAAGGCA	1260
	TGAAAGAGTA	TAATTTTATG	АААТТСАААА	TTATTCTTTT	TTCAAGTTGA	AACTIGCCIC	1320
20	TTCTACTTTA	AAAAAGTATA	TAGAACAGTT	ACTTCTAATA	ATCAGAAAGA	GATGTTTAT	1380
20	AGAACATTTC	тттаататаа	AGTTAGAGAT	GTCTTCATAG	GCAGTATGGC	TATCTTTGCC 1	1440
	ACAGAAACAT	AAGTAAAATT	TTAGAGTTCT	CTTTTCCATG	AGGTCAAAAA	TATAATTTAT	1500
25	TCCTCAGTCA	TGGTTTTCTA	ААТАТСТСТА	CTCCACATTC	CATTITAATT	GATATGAGGG	1560
	TGTTAAAGTA	CCTACTTAAT	GGGTTGATTA	CTATCAAAAT	GACCAAATTA	TACCAAAGAA	1620
20	CTTAAGAGGA	AGCACTTTCA	GAACTATTCA	CTTGCCAGGT	ATTTTCTAAA	ATTCCACCTG	1680
30	AAAGCCAAAA	GATAAAATAC	ATNAGTTGGA	TTTTAATGAT	ATAAGCATCA	CACAATTTTA	1740
	CATTAAGAAA	TACTGTGCAG	CCCATGCGTG	GTGGCTCAGG	CCTGTAATCC	CAGCANTTTG	1800
35	GGAGGCCGAG	GTGGGCAGAT	CACCGGAGGT	CAGGAGTTCG	AGACCAGCCT	TGCCAACATA	1860
	GTGAAACCCT	GTCTTTACTA	AAAATACAAA	AATTAGCCGG	GCATGGTGGC	AGGCACCTGT	1920
40	AATCCCAGCT	ACTAGGGAGG	CTTTTGAACC	CAGGAGGCAG	AGGITGCAGC	GAGCTGAGAT	1980
40	CGCGCCACTG	CACTCCAGCC	TGGGTGATAG	AGTGAGATTC	AGTCTCAAAA	AAAAAAAA	2040
	ААААААААА	AATGACCTCG	A				2061
4							

(2) INFORMATION FOR SEQ ID NO: 15:

50 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1412 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

55

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

CCCTTCATCT GCGTTGCCAG GAACCCTGTC AGCAGAAACT TCTCAAGCCC CATCCTTGCC 60

AGGAAGCTCT GTGAAGGTGC TGCTGATGAC CCAGATTCCT CCATGGTCCT CCTGTGTCTC 120

	CTGTTGGTGC	CCCTCCTGCT	CAGTCTCTT	GTACTGGGGC	TATTTCTTTC	GTTTCTGAAG	180
5 .	AGAGAGAGAC	AAGAAGAGTA	CATTGAAGAG	AAGAAGAGAG	TGGACATTTG	TCGGGAAACT	240
J	CCTAACATAT	GCCCCCATTC	TGGAGAGAAC	ACAGAGTACG	ACACAATCCC	TCACACTAAT	300
	AGAACAATCC	TAAAGGAAGA	TCCAGCAAAT	ACGGTTTACT	CCACTGTGGA	AATACCGAAA	360
10	AAGATGGAAA	ATCCCCACTC	ACTGCTCACG	ATGCCAGACA	CACCAAGGCT	ATTTGCCTAT	420
	GAGAATGTTA	TCTAGACAGC	AGTGCACTCC	CCTAAGTCTC	TGCTCAÁAAA	AAAAACAATT	480
15	CTCGGCCCAA	AGAAAACAAT	CAGAAGAATT	CACTGATTTG	ACTAGAAACA	TCAAGGAAGA	540
13	ATGAAGAACG	TTGACTTTTT	TCCAGGATAA	ATTATCTCTG	ATGCTTCTTT	AGATTTAAGA	600
			•			AATCACTTCA	660
20	TCCCAAAAAT	GGGATTGTGA	ATGTCAGCAA	ACCATAAAAA	AAGTGCTTAG	AAGTATTCCT	720
	ATAAAAATGT	AAATGCAAGG	TCACACATAT	TAATGAČAGC	CTGTTGTATT	AATGATGGCT	780
25	CCAGGTCAGT	GTCTGGAGTT	TCATTCCATC	CCAGGGCTTG	GATGTCAGGA	TTATACCAAG	840
دی	AGTCTTGCTA	CCAGGAGGGC	AAGAAGACCA	AAACAGACAG	ACAAGTCCAG	CAGAAGCAGA	900
	TGCACCTGAC	AAAAATGGAT	GTATTAATTG	GCTCTATAAA	CTATGTGCCC	AGCAYTATGC	960
30	TGAGCTTACA	CTAATTGGTC	AGACATGCTG	TCTGCCCTCA	TGAAATTGGC	TCCAAATGAW	1020
	TGAACTACTT	TCATGAGCAG	TTGTAGCAGG	CCTGACCACA	GATŢCCCAGA	GGGCCAGGTG	1080
35	TGGATCCACA	GGACTTGAAG	GTCAAAGTTC	ACAAAGATGA	AGAATCAGGG	TAGCTGACCA	1140
,,,	TGTTTGGCAG	АТАСТАТААТ	GGAGACACAG	AAGTGTGCAT	GGCCCAAGGA	CAAGGACCTC	1200
	CAGCCAGGCT	TCATTTATGC	ACTTGTCTGC	AAAAGAAAAG	TCTAGGTTTT	AAGGCTGTGC	1260
10	CAGAACCCAT	CCCAATAAAG	AGACCGAGTC	TGAAGTCACA	TTGTAAATCT	AGTGTAGGAG	1320
	ACTTGGAGTC	AGGCAGTGAG	ACTGGTGGG	CACGGGGGGC .	antgggtant	GTAAACCTTT	1380
5	TAAAGATGGT	TAATTCNICA	TTAGTGTTTT	TT			1412
_							

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1052 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

TTCCTCTCCT CTCTCTACCC CTCCTGTCTC TCCTCCCCTC CTCTCTCTCTC

60

50

	TCTCTTCCTC	TCCTCTCTCT	TCCCTTCCTG	TCTCTCTTCC	CCTCCTCTCT	CTCTTCCTGT	120
	CCTCTATCTC	TTCCCCTCCT	CTATCTCTTC	CTCTCCTCTC	TCTCTTCCTC	TCCTCTCTCT	180
5	CTCTTSCTTT	CTTCTCTCTC	TCCTGTCTCG	GCTGTTGTGG	GTTGCAGGTT	GGGTGCTGCT	240
	GTTGTGGTCC	TTCCCAGAAA	CTGCCAGTAG	AGGGCAGCCT	GGCAȚCCTĂ	ATGCTTACTC	300
10,	TGGTTGTTAC	ACAAAGAAAA	TATTGGGGTC	ACTGGCGAGC	CCACCCÁCAC	TCACCAGAAT	360
10/	CTCCACTGTA	GTCCCCCTAA	CAAACAGCCC	TICACITCCT	CTCCCACTTC	agcaätttgt	420
	ATTTTGATGC	CATTGGCCTC	AGATCAGAGT	GTTTTAAATC	ATCACGCCCT	GGCTTATCCC	480
15	TGGTCGAGCC	AGGACACGGG	GTGCTTCAGT	GGTCTGTCA	CCCTCTCTCC	TTGAAGCATG	540
	TTGCTTTTAT	TTATTTACTT	TTACTCTCAC	CCTGCTCCTG	TACCAGCAGG	GGCCACTTCA	600
20	AAGCCAAGGT	ACAGGGTGAT	AACTTGTGGT	CCAGCATCAG	TTTTCTCCAC	TTCTTTCTCC	660
20	CACTCACCCC	CAGCAAGGTG	CCTGGGGAGA	CTTGAGCAGA	TGTTTCATTT	TGGCCTGGCC '	720
	AGTGGCTGAA	AGCAGGCCTC	CAATGCACTG	TGACCTCTGG	CTTCCCCAGC	AGCTTTCCCA	780
25	GAGAGGCAGA	GGGCCTTCC	ACAGCCCGGG	TTCTCCTGCT	GCCTCCTGCC	TGCTGCAGCT	840
	GCAGGCATTC	TGAGGGGCAA	CGTGGAGGAA	GGGCCAGGGA	TGCATGGGAT	TTTAATTGTT	900
30	TCATCACACC	TTCCCCGTGG	CAAAGAAACA	GTCAGTCCTC	TTCAGGTGTC	TTCTGGATTT	960
50	CTGGTGATGG	ACAGAGAAAT	CTTTTTACAG	TTTCAAATTA	TGTTCAACAA	ATAAAAATTG	1020
	CATTTTTTAT	TTTGGAAAAA	AAAAAAAA	AA			1052

(2) INFORMATION FOR SEQ ID NO: 17:

40 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 683 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

AATTOGGCAG AGGCACTTAT CATGTACATA TAGCCTGTTT TTTAGCATTG TTAGACAAAG 60 50 TAGGCATATT CCTTTCCATC CAAGAACTCA TAACCTAGTA ATTGTAGTTG GCTGATAGCT 120 CATTGCCCAT ACACAAGGAT CTAACACAAC CTCTTGAATA AACATCCCCC TTATTCAGAA 180 ATGCCTTTTC CTATTTCCAT ATTGCAACTT TGCTTACAAA TTTCCAATCT GTCTTTCTGT 240 55 TTACAGAAGA TATACAAAAT TCCTTTTGTA TGATCTCTTT ATATCTCTTG ATTTTCTTTT 300 GTGTTTGCTA CCAAAGGGCC TGCACATAGT GAGAAGATTG TGCATGATCT GTGAGCTCTA 360 60 CCACACCTGG AATTAGGGAT CACCAATATG AGAAAAAAA TTGGAGGTAC AAATAACATT 420

	ATCATATGTW ATTOGCATAT AAATTACAGA TGTWTCTATG ACTAAAAACC CTGTGGATAT	480
5	WAACCMAATG CAGATAAWIW TAATAAAATW TWTAAAAATW TWATCMAATA ATGATAGTGC	540
_	TATTCAAATA CTTCAAATTT GCACAGTGAT TTATTTCTTA AAATATGTTA ACACATGTGA	600
	GCCAATACAC TGAGGTCACT GGATAAATAA ACAGATTCTT GCAAAAAAAA AAAAAAAAA	660
10	ACTCGAGGG GGCCCGTACC CTT	683
15	(2) INFORMATION FOR SEQ ID NO: 18:	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 1054 base pairs	
20	(B) TYPE: nucleic acid (C) STRANDEDNESS: double	
	(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:	
25	AAACTCATTT AGGTGACACT ATAGAAGGTA CGCCTGCAGG TACCGGTCCG GAATTCCCGG	60
	GTCGACCCAC GMENCCGGCG ACAAGATGGC AGCAGCGTGT CGGAGCGTGA AGGGCCTGGT	120
30	GGCGGTAATA ACCGGAGGAG CCTCGGGCCT GGGCCTGGCC ACGGCGGACG ACTTGTGGGG	180
	CAGGGAGCCT CTGCTGTGCT TCTGGACCTG CCCAACTCGG GTGGGGAGGC CCAAGCCAAG	240
	AAGTTAGGAA ACAACTGCGT TTTCGCCCCA GCCGACGTGA CCTCTGAGAA GGATGTGCAA	300
35	ACAGCTCTGG CTCTAGCAAA AGGAAAGTTT GGCCGTGTGG ATGTAGCTGT CAACTGTGCA	360
	GGCATCGCGG TGGCTAGCAA GACGTACAAC TTAAAGAAGG GCCAGACCCA TACCTTGGAA	420
40	GACTICCAGC GAGITCITGA TOTGAATCTC ATGGGCACCT TCAATGTGAT CCGCCTGGTG	480
	GCTGGTGAGA TGGGCCAGAA TGAACCAGAC CAGGGGGGCC AACGTGGGGT CATCATCAAC	540
	ACTGCCAGTG TGGCTGCCTT CGAGGGTCAG GTTGGACAAG CTGCATACTC TGCTTCCAAG	600
45	GGGGGAATAG TGGGCATGAC ACTGCCCATT GCTCGGGATC TGGCTCCCAT AGGTATCCGG	660
	GTGATGACCA TTGCCCCAGG TCTGTTTGGC ACCCCACTGC TGACCAGCCT CCCAGAGAAA	720
50	GIGIGCAACT TCTTGGCCAG CCAAGTGCCC TTCCCTAGCC GACTGGGTGA CCCTGCTGAG	780
	TATGCTCACC TCGTACAGGC CATCATCGAG AACCCATTCC TCAATGGAGA GGTCATCCGG	840
	CTGGATGGGG CCATTCGTAT GCAGCCTTGA AGGGAGAAGG CAGAGAAAAC ACACGCTCCT	900
55	CTGCCCTTCC TTTCCCTGGG GTACTACTCT CCAGCTTGGG AGGAAGCCCA GTAGCCATTT	960
	TGTAACTGCC TACCAGTCGC CCTCTGTGCC TAATAAAGTC TCTTTTTCTC ACANAAAAAA	1020
60	AAAA AAAAAAAA AAAAAAAA AAAAAAAAA	1054

(2) INFORMATION FOR SEQ ID NO: 19:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1393 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

10

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:	
15	GGAACAAGCT GGGATATGTG AGCGTTAAGC TACTCACATC CTTCAAAAAG GTGAAACATC	60
13	TTACACGGGA CTGGAGAACC ACAGCACATG CTTGAAGTA TTCAGTGGTC CTTGAGTTGA	120
	ATGAGGNCCA CCGGAAGGTG AGGAGGACCA CCCCCGTCCC ACTGTTCCCC AACGAGAACC	180
20	TCCCCAGCAA GATGCTCCTG GTCTATGATC TCTACTTGTY TCCTAAGCTG TGGGCTCTGG	240
	CCACCCCCA GAAGAATGGG AAGGGTGCAA GARAAGGTGA TGGAACACCT GCTCAAGCTT	300
25	TTTGGGACTT TTGGAGTCAT CTCATCAGTG CGGATCCTCA AACCTGGGAG AGAGCTGCCC	360
	CCTGACATCC GGAGGNTCCA GCAGCCGCTA CAGCTCCTCT GACCCCGAGA GCAACCCCAC	420
	ATCCCCTATG GCGGGCCGAC GGCACGNGKC CACCAACAAG CTCAGCCCGT CTGGCCACCA	480
30	GAATCTCTTT CTGAGTCCAA ATGCCTCCCC GTGCACAAGT CCTTGGAGCA GCCCCTTGGC	540
	CCAACGCAAA GGCGTTTCCA GAAAGTCCCC ACTGGCGGAG GAAGGTAGAC TGAACTGCAG	600
35	CACCAGCCCT GAGATCTTCC GCAAGTGTAT GGATTATTCC TCTGACAGCA GCGTCACTCC	660
	CTCTGGCAGC CCCTGGGTCC GGAGGCGTCG CCAAGCCGAG ATGGGGGACCC AGGAGAAAAG	720
	CCCCGGTACG AGTCCCCTGC TCTCCCGGAA GATGCAGACT GCAGATGGGS TACCCGTAGG	780
40	TNGCTTGAGG TTGCCCAGGG GTCCTGACAA CACCAGAGGA TTTCATGGCC ATGAGAGGAG	840
	CAGGGCCTGT GTATAAATAC CTTCTATTTT TAATACAAGC TCCACTGAAA ACCACCTTCG	900
45	TTTTCAAGGT TCTGACAAAC ACCTGGCATG ACAGAATGGA ATTCGTTCCC CTTTGAGAGA	960
	TTTTTTATTC ATGTAGACCT CTTAATTTAT CTATCTGTAA TATACATAAA TCGGTACGCC	1020
	ATGGTTTGAA GACCACCTTC TAGTTCAGGA CTCCTGTTCT TCCCAGCATG GCCACTATTT	1080
50	TGATGATGGC TGATGTGTGT GAGTGTGATG GCCCTGAAGG GCTGTAGGAC GGAGGTTCCC	1140
	TGGGGGAAGT CTGTTCTTTG GTATGGAATT TTTCTCTCTT CTTTGGTATG GAATTTTTCC	1200
55	CTTCAGTGAC TGAGCTGTCC TCGATAGGCC ATGCAAGGGC TTCCTGAGAG TTCAGGAAAG	1260
	TTCTCTTGTG CAACAGCAAG TAGCTAAGCC TATAGCATGG TGTCTTGTAG GACCAAATCG	1320
	ATGTTACCTG TCAAGTAAAT AAATAATAAA ACACCCAACT GGGAGTGCTG AAAAAAAAANA	1380
60	ANNAAAAAC TCG	1393

_							
5	(2)	INFORMATION	FOR	SEQ	ID	NO:	20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1215 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

	-	_			,		
15	AGGAAAAGTT	TTCCNAATTG	GAAAGCGGGC	AGTGAGCGCA	ACGCA ATTAA	TGTGAGTTAG	60
	NTCANTCATT	AGGCACCCCA	GGCTTTACAC	TTTATGCTTC	CGGNTCGTAT	GTTGTGTGGA	120
20	ATTGTGAGCG	GATAACAATT	TCACACAGGA	AACAGCTATG	ACCATGATTA	CGCCAAGCTN	180
	TAATACGACT	CACTATAGGG	AAAGCTGGTA	CGCCTGCAGG	TACCGGTCCG	GAATTCCCGG	240
	GTCGACCCAC	GCGTCCGCCC	ACGCGTCCGT	GAAAATCCGA	AGTGCCGCGG	AAAGTGGAGG	300
25	TGAGGGCCGC	CCGCCCTAGA	GGTGCCCGTC	CGAGAGGCAG	AGCTGACAAG	GAAGGTTTCG	360
	AGCGTTTTGC	TGGCAAAGGG	ATTTCTTACA	ACCTCCAGGC	ATGCGTCTTT	CTGCCCTGCT	420
30	GGCCTTGGCA	TCCAAGGTCA	CTCTGCCCCC	CCATTACCGC	TATGGGATGA	GCCCCCCAGG	480
	CTCTGTTGCA	GACAAGAGGA	AGAACCCCCC	ATGGATCAGG	CGGCGCCCAG	TGGTTGTGGA	540
	ACCCATCTCT	GATGAAGACT	GGTATCTGTT	CTGTGGGGAC	ACGGTGGAGA	TCCTAGAAGG	600
35	CAAGGATGCC	GGGAAGCAGG	GCAAAGTGGT	TCAAGTTATC	CGGCAGCGAA	ACTGGGTGGT	660
	CCTCCCACCC	CTGAACACAC	ATTACCGCTA	CATTGGCAAG	ACCATGGATT	ACCGGGGAAC	720
40	CATGATCCCT	AGTGAAGCCC	CCTTGCTCCA	CCGCCAGGTC	AAACTTGTGG	ATCCTATGGA	780
	CAGGAAACCC	ACTGAGATCG	ACTGGAGATT	TACTGAAGCA	GGAGAGCGGG	TACGAGTCTC	840
	CACACGATCA	GGGAGAATTA	TCCCTAAACC	CGAATTTCCC	AGAGCTGATG	GCATCGTCCC	900
45	TGAAACGTGG	ATTGATGGCC	CCAAAGACAC	ATCAGTGGAA	GATGCTTTAG	AAAGAACCTA	960
	Terecetet	CTAAAGACAC	TGCAGGAGGA	GGTGATGGAG	GCCATGGGGA	TCAAGGAGAC	1020
50	CCGGAAATAC	AAGAAGGTCT	ATTGGTATTG	AGCCTGGGGC	AGAGCAGCTC	CTCCCCAACT	1080
	TCTGTCCCAG	CCTTGAAGGC	TGAGGCACTT	CTTTTTCAGA	TGCCAATAAA	GAGCACTTTA	1140
	TGAGTCCTCC	AAAAAAAA	АААААААА	АААААААА	АААААААА	AAAAAAAA	1200
55	AAAAGGGGCG	eccec					1215

^{60 (2)} INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2042 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

				·			
10	CTGCATCCAG	GCGCAGAATA	ACCTGGGTAT	CTTGTGGTCT	GAAAGAGAGA	AATTGAAACT	60
	GCACAGGCTT	ACCTAGAGTC	ATCAGAAGCA	CTATATAATC	AGTATATGAA	AGAGGTTGGG	120
15	AGTCCTCCTC	TTGATCCTAC	TGAGCGTTTT	CTTCTGAAGA	AGAGAAACTT	ACTGAACAAG	180
15	AGAGATCAAA	AAGATTTGAA	AAGGTTTATA	CTCATAACCT	ATATTACCTA	GCTCAAGTCT	240
	ACCAGCATCT	GGAAATGTTT	GAGAAGGCTG	CTCACTATTG	CCATAGTACA	CTAAAACGCC	300
20	AGCTTGAGCA	CAATGCCTAC	CATCCTATAG	AGTGGGCTAT	CAATGCTGCT	ACCTTGTCAC	360
	ACTITTACAT	CAATAAGCTA	TGCTTTATGG	AGGCCAGGCA	CTGTTTATCA	GCTGCTAATG	420
25	TCATTTTTGG	TCAAACTGGA	AAGATCTCAG	CCACAGAAGA	CACTCCTGAA	GCTGAAGGAG	480
25	AAGTGCCAGA	GCTTTATCAT	CAAAGAAAGG	GGGAAATAGC	AAGGTGCTOG	ATCAAATACT	540
	GTTTGACTCT	CATGCAGAAT	GCCCAACTCT	CCATGCAGGA	CAACATAGGA	GAGCTTGATC	600
30	TTGATAAACA	GTCTGAACTT	AGAGCTTTAA	GGAAAAAAGA	ACTAGATGAG	GAGGAAAGCA	660
	TTCGGAAAAA	AGCTGTGCAG	TTTGGAACCG	GTGAACTGTG	TGATGCCATC	TCTGCAGTAG	720
35	AAGAGAAAGT	GAGCTACTTG	AGACCTTTAG	ATTITGAAGA	AGCCAGAGAA	CTTTTCTTAT	['] 780
<i>J J</i>	TGGGTCAGCA	CTATGTCTTT	GAGGCAAAAG	AGTTCTTTCA	GATTGATGGT	TATGTCACTG	840
	ACCATATTGA	AGITGTCCAA	GACCACAGTG	CTCTGTTTAA	GGTGCTTGCA	TTCTTTGAAA	900
40	CTGACATGGA	GAGACGGTGC	AAGATGCATA	AACGCRGAAT	AGCCATGCTA	GAGCCCCTAA	960
	CTGTAGACCT	GAATCCACAG	TATTATCTGT	TGGTCAACAG	ACAGATCCAG	TTTGAAATTG	1020
45	CACATGCTTA	CTATGATATG	ATGGATTTGA	AGGTTGCCAT	TGCTGACAGG	CTAAGGGATC	1080
	CTGATTCACA	CATTGTAAAA	ATAAATA	ATCTTAATAA	GTCAGCACTG	AAGTACTACC	1140
	AGCTCTTCTT	AGACTCCCTG	AGAGACCCAA	ATAAAGTATT	CCCTGAGCAT	ATAGGGGAAG	1200
50	ATGTTCTTCG	CCCTGCCATG	TTAGCTAAGT	TTCGAGTTGC	CCGTCTCTAT	GGCAAAATCA	1260
	TTACTGCAGA	TCCCAAGAAA	GAGCTGGAAA	ATTTGGCAAC	ATCATTGGGA	ACATTACAAA	1320
55	TTTATTGTTG	ATTACTGTGA	AAAGCATCCT	GAGGCCGCCC	AGGAAATAGA	AGTTGAGCTA	1380
	GAACTTAGTA	AAGAGATGGT	TAGTCTTCTC	CCAACAAAAA	TGGAGAGATT	CAGAACCAAG	1440
	ATGGCCCTGA	CTTAATCCTT	GTTTTTAAAG	AAAGGAAATG	TGCAATATTG	AAGTGATCTT	1500
60	TTTCCCTAGT	CAGACAGGCC	CAATTCCATT	GIGATGTITA	CCTTTATAGC	CAGGTGAGTG	1560

00					•		
	TT	t					2042
	TGTTAGAAAT	AÄAATAAACT	GACTTATTTC	ACTAATGAAA	AAAAAAAA	АААААААА	2040
15	CTAAGTTATG	TTATTATAAA	GTGTAAAATG	GTTTGTCTTA	ATTATAGGAG	AAAAAGGCCT	1980
	AATATTTGAC	TTCCTACATT	CCCCCACCC	AAAATCTTTC	CCTTTTGAAA	АТАСТАААА	1920
10	GTAACAAGCC	TGTŢAATAŢA	TTAAGATTGA	AAAAGTAACT	TCTATAGTTA	CTCCTTCTAA	1860
	TATTTGTTGG	CTAGTACTTG	ATAGATTCCT	TGTAAGAAAA	AATGCTGGGT	AATGTACCTG	1800
	TTAGATGCTT	GTTTCCTATT	AAAATACAGA	CATTTCTACC	CTCAGTTTCT	AAATGTAGAC	1740
5 .	TAATTAAAA	CTTACACCTA	ATTATGTAAA	TIGCCTIGIT	AAAGACATGT	GATTTGTATT	1680
	CAGTTTGAAC	TTGAGATACA	GTCAACTGAG	TGTTTGCTAG	GATCCTAAGG	AACATAAAGT	1620

(2) INFORMATION FOR SEQ ID NO: 22:

25 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1872 base pairs
- (B) TYPE: nucleic acid (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

GGGTCGACCC ACGCGTCCGA TTGGCCTAGA GCTCCTGTGA CCGAGAGCGC CACGGAAGCC

					•		
35	TGGGGATGAT	GTCGGGCAGC	TTTATTCTTT	GCTTGGCTTT	GGTAACTAGG	TGGTCCCCTC	120
	AAGCATCCTC	AGTICCTCTT	GCTGTTTATG	AATCTAAGAC	AAGGAAGTCC	TATAGAAGCC	180
40	AAAGGGACAG	GGACGGAAAG	GACAGGTCCC	AAGGGATGGG	GCTGTCTTTA	CTTGTGGAAA	240
	CCAGGAAATT	GCTCCTCTCA	GCCAACCAAG	GTTGACCACA	CACCACCCTT	CCGGAGCAGC	300
	TCAGTCAGCC	CTCGGGGACG	RGAAACCACA	AGCGCAGAGA	CGCTGAGGCC	CAGGCAGGTG	360
45	AAGAGGAAGT	GGCTTTGGGT	TTTTAAAGTA	GGTGAGCGTG	ACCTCTCTGA	CIGCITCITC	420
	CCCGGGGGG	ACTGCAAACC	GCTCAGGGTT	GCGGCAGAGC	CATGGACTTC	CGGTCCCTGC	480
50	AACGGGTGAC	CTAAGCGTGG	TGCACCCATC	AGTCACGCAG	GAGGACTGAC	TTGACAGACG	540
	AAAGACAAGC	CCGGATGACA	CAGGGTGAGA	AGAGTCAGGG	CCGCACCTCT	GTCCCTGCAA	600
	ACCAACAGGT	GCATGGTGAG	TGTGGCAGTC	CCCACAGCTC	CACAATGGGC	TCCCCCGCCA	660
55	ACGGGGACGA	CAGGGATCTT	CAGGAACTTC	TGACCTCACC	AAGTCAAGTG	GACCACTCTC	720
	CACTCCACGA	GGATGTGAAA	CGGTTCTTTA	AAATGGGATT	TTAGAGCCTC	GGGAATGCAT	780
60	GTGCGTCGCA	TCTTTCATAT	TATGGGTCAG	GATAGATTCA	TTTCTTGCAA	CATAGTGGAA	840

					•		
	AAGATATAAG	CTGCAGTAAT	TTGCTCTTTG	AATGACCGTC	ACCCCCAGTA	TAGGATATGC	900
•	TTGTATCCCC	CCGTCACTCC	TCCGCCTGTT	TTTTAAACTT	TTCCACCACC	TGCGTCCAAA	960
5	AAGAATGTTA	TAGCGAGTGC	TCTTAAATGT	TGAACCTGGG	TGTTGCTTCC	GGGCCAGTCT	1020
	GCGTGGCTCC	ATGAAAAGCT	CACTGCTGCC	CCAGCCGGGC	TTCTTAGAGG	AGGTCAGTTG	1080
10	TCCTATGTAT	CATCATTTAC	TCTGGGAATC	CTACTGTGAA	ATCATGTCTG	TATTITTCTG	1140
10	GAGCAGTTCA	CATAGAGTAG	AATGTGGAAT	TTCCCGTGAA	CGTCTCCTTC	CTCCCCCGTA	1200
	TCTGCCGCCT	GTCACTTCGC	CACCGTGCTA	GAATACTGTT	GŢGTTGTAAG	ATGACTAATT	, 1260
15	TTAAAAGAAC	CTCCCTGAA	AAGTTCTTAG	AAACGCAATG	AAAGGGAGGA	ACTTGTCCTT	1320
	TACCCAGTTT	TTCCTTTGTA	GGATGGGAAA	GTATAAAAAG	GCACAGAAGG	TTGTCATGGG	1380
20	CTGTTCCTTG	GGGGTTTTTA	TCCTGCTCAC	CGTGGAGATA	AGCCTGCGGC	TTGTCTAACC	1440
20	AGCGCAGCGM	AAAGGTCTCA	ATGCCTTTTG	GTAACATCCG	TCATTGCAGA	AGAAAGTTTA	1500
	CACGACGTCA	AAAAGTGACG	TTCATGCTAA	GTGTTTTTCC	AGAAATATTG	GTTTCATGTT	1560
25	TCTTATTKGC	TCTGCCTCCT	GTGCTTATAT	CATCCAAAAA	СТТТТТАААА	AGGTCCAGAA	1620
	TTCTATTTTA	ACCTGATGTT	GAGCACCTTT	AAAACGTTCG	TATGTGTGTT	GCACTAATTC	1680
30	TAAACTTTGG	AGGCATITIG	CTGTGTGAGG	CCGATCGCCA	CTGTAAAGGT	CCTAGAGTTG	1740
50	CCTGTTTGTC	TCTGGAGATG	GAATTAAACC	AAATAAAGAG	CTTCCACTGG	AGGCTTGTAT	1800
	TGACCTTGTA	ACTATATGTT	AATCTCGTGT	ТААААТАААА	TATAACTTGT	GAAAAAAAA	1860
35	ААААААААА	NT					1872

40 (2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 289 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

50	CATTTACCCA	CCTATCAACA	TGTTTGCTTT	CICITITGIT	GGTGAGAATG	AGTGGCTTCT	60
55	TGCTCCTAGC	TAGAGCCAGT	CCTTCCATAT	GTGCTTTAGA	TTCTTCCTGT	TTTGTTCAAG	120
	AATATTGCTC	AAGCTATTCT	TCCTCCTGTT	TCCTGCATCA	GCATTTCCCC	TCTCTACTAG	180
	ATCATCTCTG	TCAGTAAATG	AACATGTTGT	TGTTTCTCCT	AGAAGTACTG	TTTCTATATC	240
	TAGATAGTAC	TCTAGCTAGA	GTTAAAAAAA	ааааааааа	CCTNGGGGG		289

(2) INFORMATION FOR SEQ ID NO: 24:

5	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 3533 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	•
10	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:	
	TITTATITAC TICAAATTAA CIGIACITTA CICAAATAGA AAANGAATAA TITICACATT	60
15	ATGAAGCTAC ACAATTCCAA AATACACATG CTGAGGCTCT TTTTAAGTCC GAATTGTCTA.	120
	GTAATTACAA AAAAGTGAAG AGTTTACAGA TATACAAGGA AATAAAGGCG AATTATTGCA	180
20	AAGAAAACAA GITTAATTIC ACTITGAATG ACAACGATTI TICTGGAAAG CAGATACTIC	240
	ACTCCTTTAA GTTTCCACCC AAGCCACAAT AATTTCAAAC GGTCTTGCGG ATGACCCAGC	300
	TGGTCACTCT TGTTTATGTG GGGACTGGAG GTAATGAGAG CCAAAAAAAG TGCTATAAAC	360
25	CTAATTIGGC TAGAGCAAGT TCACACGACA CGACCGIGCT TTAAAAACTT GCTCTCCATT	420
	ATGUACTUCC TUCCATCAGG TUGGGGAAAA AAAAATGGUG GGGATGGUGA GUAAACACAC	480
30	CAGTGGTTTC ATCAGAGGG AACTCACTAC TCAGGAGGTG ACGGTGACGT GGTGCCGGTC	540
	CCTGAAGTAC GCGCACAAGC TCCGGAGGTT GCGGGAGGTT CCGCTGCCGC CTGGAGGGAA	600
	GCCGGAGCGA CGGGGGTCAC GGCGGCGGTC AGAGGGTAAA GGTCTTGCTC CCAGCAGCCT	660
35	CCGCGGTGGA TACGTCGCCA TCTTGGATCC GCGGGACAAG AAAATTCATG CGAGGGAGAC	720
	GTGGTGGGCG GTCCTTCCTG TGACACGACC CTTGAGTGAC AGTTCTATTT GATTGCCTCC	780
40	GGTACTGTGA GGAAAGGACA CGACTCTATG GTGAGGACTG ATGGACATAC ATTATCTGAG	840
	AAAAGAAACT ACCAGGTGAC AAACAGCATG TITGGTGCTT CAAGAAAGAA GTTTGTAGAG	900
	GGGGTCGACA GTGACTACCA TGACGAAAAC ATGTACTACA GCCAGTCTTC TATGTTTCCA	960
45	CATCGGTCAG AAAAAGATAT GCTGGCATCA CCATCTACAT CAGGTCAGCT GTCTCAGTTT	1020
	GGGGCAAGTT TATACGGGCA ACAAAGTGCA CTAGGCCTTC CAATGAGGGG GATGAGCAAC	1080
50	AATACCCCTC AGITAAATCG CAGCTTATCA CAAGGCACTC AGITACCGAG CCACGTCACG	1140
	CCAACAACAG GGGTACCAAC AATGTCACTT CACACGCCTC CATCTCCAAG CAGGGGTATT	1200
	TTGCCTATGA ATCCTARGAA TATGATGAAC CACTCCCAGG TTGGTCAGGG CATTGGAATT	1260
55	CCTAGCAGGA CAAATAGCAT GAGCAGTTCA GGGTTAGGTA GCCCCAACAG AAGCTCGCCA	1320
	AGCATAATAT GTATGCCAAA GCAGCAGCCT TCTCGACAGC CITITACTGT GAACAGTATG	1380
60	TCTGGATTTG GAATGAACAG GAATCAGGCA TTTGGAATGA ATAACTCCTT ATCAAGTAAC	1440

	ATTTTTAATG	GAACAGACGG	AAGTGAAAAT	GTGACAGGAT	TGGACCTTTC	AGATITICCCA	1500
,	GCATTAGCAG	ACCGAAACAG	GAGGGAAGGA	AGTGGTAACC	CAACTCCATT	AATAAACCCC	1560
5	TTGGCTGGAA	GAGCTCCTTA	TGTTGGAATG	GTAACAAAAC	CAGCAAATGA	ACAATCCCAG	1620
	GACTTCTCAA	TACACAATGA	AGATTTTCCA	GCATTACCAG	GCTCCAGCTA	TAAAGATCCA	1680
10	ACATCAAGTA	ATGATGACAG	TAAATCTAAT	TTGAATACAT	CTGGCAAGAC	AACTTCAAGT	1740
	ACAGATGGAC	CCAAATTCCC	TGGAGATAAA	AGTTCAACAA	CACAAAATAA	TAACCAGCAG	1800
	AAAAAAGGGA	TCCAGGTGTT	ACCTGATGGT	CGGGTTACTA	AÇATTCCTCA	AGGGATGGTG	1860
15	ACGGACCAAT	TTGGAATGAT	TGGCCTGTTA	ACATTTATCA	GGCAGCAGA	GACAGACCCA	1920
	GGAATGGTAC	ATCTTGCATT	AGGAAGTGAC	TTAACAACAT	TAGGCCTCAA	TCTGAACTCT	1980
20	CCTGAAAATC	TCTACCCCAA	ATTIGCGICA	CCCTGGGCAT	CTTCACCTTG	TCGACCTCAA	2040
	GACATAGACT	TCCATGTTCC	ATCTGAGTAC	TTAACGAACA	TTCACATTAG	GGATAAGCTG	2100
	GCTGCAATAA	AACTTGGCCG	ATATGGTGAA	GACCTTCTCT	TCTATCTCTA	TTACATGAAT	2160
25	GGAGGAGACG	TATTACAACT	TTTAGCTGCA	GTGGAÇCTTT	TTAACCGTGA	TTGGAGATAC	2220
	CACAAAGAAG	AACGAGTATG	GATTACCAGG	GCACCAGGCA	TGGAGCCAAC	AATGAAAACC	2280
30	AATACCTATG	AGAGGGGAAC	ATATTACTTC	TTTGACTGTC	TTAACTGGAG	GAAAGTAGCT	2340
	AAGGAGTTCC	ATCTGGAATA	TGACAAATTA	GAAGAACGGC	CTCACCTGCC	ATCCACCTTC	2400
	AACTACAACC	CTGCTCAGCA	AGCCTTCTAA	АААААААА	AAAAAAAAA	AAAAAGACTT	2460
35	CCCTTTTCTT	GGGGTATGGC	TGTCTCAGCA	CAATACTCAA	CATAACTGCA	GAACTGATGT	2520
	GGCTCAGGCA	CCCTGGTTTT	AATTCCTTGA	GGATCTGGCA	ATTGGCTTAC	GCAAAAGGTC	2580
40	ACCATTIGAG	GTCCTGCCTT	ACTAATTATG	TGCTGCCCAA	CAACTAAATT	TGTAATTTGT	2640
	TTTTCTCTAG	TTTGAGCAGG	GTCTGAATTT	TTTCATTTAT	TTCCTTTTTT	GCCAGCAGAC	2700
	AGACTTGAGT	CTGTAAAGAC	AAGCAAATAC	ACTGACAGAA	GTTTACCATA	GTTTCTAAAA	2760
45	TGTAAAAAAG	AAAACCCCCA	AAAGACTCAA	GAAAATTAGA	CCACAAATTT	TGCATTGTTC	2820
	ATTGTAGCAC	TATTGGTAAT	AAAATAACAA	ATGTTTGTGC	ATTTTTATGT	GAAGATCCTT	2880
50	CTCGTATTTC	ATTTGGAAAG	ATGAGCAAGA	GCTCTGCTTC	CTTCATTTTA	CTTCCCCTTC	2940
	TGTTTTTGAA	AGGCAGTTTC	GCCAAGCTTA	ATGCAAGAAT	ATCTGACTGT	TTAGAAGAAA	3000
	GATATTGCCA	CAATCTCTGG	ATGGTTTTCC	AGGGTTGTGT	TATTACTGAG	CTTCATCTTT	3060
55	CCAGAATGAG	CAAAACACTG	TCCAGTCTTT	GTTACGATTT	TGTAATAAAT	GTGTACATTT	3120
	TTTTTAAATT	TTTGGACATC	ACATGAATAA	AGGTATGTAT	GTACGAATGT	GTATATATTA	3180
50	TATATATGAC	ATCTATTTG	GAAAATGTTT	CCCTCCTCT	ACCTCATTTT	TAGGAGGTGT	3240

	GCATGGATGC	AATATATGAA	AATGGGACAT	TCTGGAACTG	CTGGTCAGGG	GACTITGICG	3300
	CCCTGTGCAC	TAAAAGGGCC	AGATTTTCAG	CAGCCAAGGA	CATCCATACC	CAAGTGAATG	3360
5	TGATGGGACT	TAAAAGAAGT	GAACTGAGAC	AATTCACTCT	GCTGTTTGA	ACAGCAGCGT	3420
	TTCATAGGAA	GAGAAAAAA	GATCAATCTT	GTATTTTCTG	ACCACATAAA	GGCTTCTTCT	3480
10	CTTTGTAATA	AAGTAGAAAA	GCTCTCCTCA	ĄAAAAAAA	AAAAAAACTC	GAG	3533

(2) INFORMATION FOR SEQ ID NO: 25:

15

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1148 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

20

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

25	ACCCACGCGT	CCGCAAATTA	TACTTCCTCA	TŢCATATTAT	GTTGATACAA	AAGACCTTGG	60
20	CAGCCATTTC	TCCCAGCAGT	TTTAAAGGAT	GAACATTGGA	TTTCATGCCA	TCCCATAGAA	120
	AACCTGTTTT	AAAATTITAG	GGATCTTTAC	TTGGTCATAC	ATGAAAAGTA	CACTGCTTAG	180
30	AAATTATAGA	CTATTATGAT	CTGTCCACAG	TGCCCATTGT	CACTTCTTTG	TCTCATTTCT	240
	TCCCTTTCTT	CCTTAGTCAT	CCAAATAAGC	CTGAAAACCA	TAAGAGATAT	TACTTTATTG	300
35	AATATGGTTG	GCATTAAATT	TAGCATTTCA	TTATCTAACA	AAATTAATAT	AAATTCCAGG	, 360
33	ACATGGTAAA	ATGTGTTTTA	ATAACCCCCA	GACCCAAATG	AAAATTTCAA	AGTCAATACC	420
	AGCAGATTCA	TGAAAGTAAA	TTTAGTCCTA	TAATTTTCAG	CTTAATTATA	AACAAAGGAA	480
40	CAAATAAGTG	GAAGGGCAGC	TATTACCATT	CGCTTAGTCA	AAACATTCGG	TTACTGCCCT	540
	TTAATACACT	CCTATCATCA	GCACTTCCAC	CATGTATTAC	AAGTCTTGAC	CCATCCCTGT	600
45	CGTAACTCCA	GTAAAAGTTA	CTGTTACTAG	AAAATTTTTA	TCAATTAACT	GACAAATAGT	660
43	TTCTTTTTAA	AGTAGTTTCT	TCCATCTTTA	TTCTGACTAG	CTTCCAAAAT	GTGTTCCCTT	720
	TTTGAATCGA	GGTTTTTTTG	TTTTGTTTTG	TTTTCTGAAA	AAATCATACA	ACTITGIGCT	780
50	TCTATTGCTT	TTTTGTGTTT	TGTTAAGCAT	GTCCCTTGGC	CCAAATGGAA	GAGGAAATGT	840
	TTAATTAATG	CTTTTTAGTT	TAAATAAATT	GAATCATTTA	TAATAATCAG	TGTTAACAAT	900
55	TTAGTGACCC	TTGGTAGGTT	AAAGGTTGCA	TTATTTATAC	TTGAGATTTT	TTTCCCCTAA	960
33	CTATTCTGTT	TITIGTACIT	TAAAACTATG	GGGGAAATAT	CACTGGTCTG	TCAAGAAACA	1020
	GCAGTAATTA	TTACTGAGTT	AAATTGAAAA	GTCCAGTGGA	CCAGGCATTT	СТТАТАТААА	1080
60	TAAAATTGGT	GGTACTAATG	TGAAAAAAA	АААААААА	AACTCGAGGG	GGGCCCGGTA	1140

WO 98/42738 PCT/US98/05311

	·	
	CCCTATTA	1148
5	•	•
	(2) INFORMATION FOR SEQ ID NO: 26:	
10	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 717 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
15	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:	
	GGCACGAGCT AGCTGCCGCC ACCCGAACAG CCTGTCCTGG TGCCCCGGCT CCCTGCCCCG	60
20	CGCCCAGTCA TGACCCTGCG CCCCTCACTC CTCCCGCTCC ATCTGCTGCT GCTGCTGCTG	120
20	CTCAGTGCGG CGGTGTGCCG GGCTGAGGCT GGGCTCGAAA CCGAAAGTCC CGTCCGGACC 1	180
	CTCCAAGTGG AGACCCTGGT GGAGCCCCCA GAACCATGTG CCGAGCCCGC TGCTTTTGGA	240
25	GACACGCTTC ACATACACTÀ CACGGGAAGC TTGGTAGATG GACGTATTAT TGACACCTCC	300
	CTGACCAGAG ACCCTCTGGT TATAGAACTT GGCCAAAAGC AGGTGATTCC AGGTCTGGAG	360
20	CAGAGTCTTC TCGACATGTG TGTGGGAGAG AAGCGAAGGG CAATCATTCC TTCTCACTTG	420
30	GCCTATGGAA AACGGGGATT TCCACCATCT GTCCCAGCGG ATGCAGTGGT GCAGTATGAC	480
	GTGGAGCTGA TTGCACTAAT CCGAGCCAAC TACTGGCTAA AGCTGGTGAA GGGCATTTTG	, 540
35	CCTCTGGTAG GGATGGCCAT GGTGCCAGCC CTCCTGGGCC TCATTGGGTA TCACCTATAC	600
	AGAAAGGCCA ATAGACCCAA AGTCTCCAAA AAGAAGCTCA AGGAAGAGAA ACGAAACAAG	660
40	AGCAAAAGA AATAATAAAT AATAAATTTT AAAAAAAAA AAAAAA	717
45	(2) INFORMATION FOR SEQ ID NO: 27: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1099 base pairs	
50	(B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:	
55	GGCACGAGCC GATGTGGACA TCATCCTGTC TATCCCCATG TTCCTGCGCC TGTACCTGAT	60
<i>J</i> .J	CGCCCGAGTC ATGCTGCTGC ACAGCAAGCT CTTCACCGAT GCCTCGTCCC GCAGCATCGG	120
	GGCCCTCAAC AAGATCAACT TCAACACCCG CTTTGTCATG AAGACGCTCA TGACCATCTG	186
60	COMPARED CHARACTER MAINTANCEAN CHARACTER AND AND ANTACTED CHARACTER	24

	CCGTGTCTGT	GAAAGTCCTG	AATCACCAGC	CCAGCCTTCT	GCTCATCAC	TTCCTGCTTG	300
5	GTACCATGAC	CAGCAGGACG	TAACTAGTAA	CTTTCTGGGT	GCCATGTGGC	TCATCTCCAT	360
	CACATTCCTT	TCCATTGGTT	ATGGGGACAT	GGTGCCCCAC	ACATACTGTG	GGAAAGGTGT	420
	CTGTCTCCTC	ACTGGCATCA	TGGGTGCAGG	CTGCACTGCC	CTTCTCCTCC	CCGTGGTGGC	480
10	CCGAÄAGCTG	GAACTCACCA	AAGCGGAGAA	GCACGTTCAT	AACTTCATGA	TGGACACTCA	540
	GCTCACCAAG	CGGATCAAGA	ATGCTGCAGC	CAATGTCCTT	CGGGAAACAT	GGTTAATCTA	600
15	TAAACACACA	AAGCTGCTAA	AGAĄGATTGA	CCATGCCAAA	GTGAGGAAAC	ACCAGAGGAA	660
	GTTCCTCCCA	AGCTATCCAC	CAGTTTGAGG	AGCGTCCCAG	ATGGAACAGA	GGAAAGCTGA	720
	GTGACCAAGC	CAACACTCTG	GIGGACCITI	CCAAGATGCA	GAATGTCATG	TATGACTTAA	780
20	TCACAGAACT	CAATGACCGG	AGCGAAGACC	TGGAGAAGCA	GATTGGCAGC	CTGGAGTCGA	840
	AGCTGGAGCA	TCTCACCGCC	AGCTTCAACT	CCCTGCCGCT	CCTCATCGCC	GACACCCTGC	900
25	GCCAGCAGCA	GCAGCAGCTC	CTGTCTGCCA	TCATCGAGGC	CCGGGGTGTC	AGCGTGGCAG	960
	TGGGCACCAC	CCACACCCCA	ATCTCCGATA	GCCCCATTGG	GGTCAGCTCC	ACCICCTICC	1020
	CGACCCCGTN	CACAAGTTCA	AGCAGTTGCT	AAATAAATCT	CCCCACTCCA	GAAGCATTAA	1080
30	AAAAAAAA	ААААААА					1099

35 (2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 941 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

45	AATTCGGCAG AGAGCCAACC GAGGGCGTTC CTGTCGGGGC TGCAGCGGCG GGAGGGAGCC	60
	CAGTGGAGGC GCCCTCCCGA AGCGCCACTG CCCATGCTGA CCACCCAGCC CTCCGGCTGC	120
50	TGATGTCATG AGTAACACCA CTGTGCCCAA TGCCCCCCAG GCCAACAGCG ACTCCATGGT	180
	GGGCTATGTG TTGGGGGCCCT TCTTCCTCAT CACCCTGGTC GGGGTGGTGG TGGCTGTGGT	240
	AATGTATGTA CAGAAGAAAA AGCGGGTGGA CCGGCTGCGC CATCACCTGC TCCCCATGTA	300 -
55	CAGCTATGAC CCAGCTGAGG AACTGCATGA GGCTGAGCAG GAGCTGCTCT CTGACATGGG	360
	AGACCCCAAG GTGGTACATG GCTGGCAGAG TGGCTACCAG CACAAGCGGA TGCCACTGCT	420
60	GGATGTCAAG ACGTGACCTG ACCCCCTTGC CCCACCCTTC AGAGCCTGGG GTYCTGGACT	480

					•		
	GCCTGGGGCC	CTGCCATCTG	CTTCCCCTGC	TGTCACCTGG	steececiec	TEGETECTEG	540
	GTCTCCATTT	CTCCCTCCAC	CCACCCTCAG	CAGCATCTGC	TTCCCATGCC	CTCACCATCA	600
5	CCTCACTGCC	CCCAGGCCTT	CTGCCCTTTG	TGGGTGTTGA	GCTCACCGCC	CACCCACAGG	660
	CACTCATGGG .	AAGAGGCTTT	CCTTCTGGGA	TGCCGCCGCC	TGGTAGACAC	CTTTGCTTTC	720
10	TCTAGCCCTC	CIGGCIGGG	CTTGGGCACA	AATCCCCAGG	CAGGCTTTGG	AGTTGTTTCC	780
10	ATGGTGATGG	GGCCAGATGT	ATAGTATTCA	GTATATATTT	TGTAAATÁAA	ATGTTTTGTG	840
	GCTAAAAAAA	AAAAAAAA	ATCNAAGGGG	GGGCCGGTAC	CCAAATTCCC	CCTATANTGA	900
15	ATTCGTATTA	ACAATTCACT	TGGGGCCGTC	CTTTTAANAA	C		941
		·		Makab se maata			
20	(2) INFORMA	TION FOR SE	O ID 100: 29) <u>.</u>			
	(2) ====================================			· ·			
25	(i)	(B) TYPI (C) STRI	MARACTERIST FTH: 756 bas E: nucleic a ANDEDNESS: O DLOGY: line	se pairs acid double		A .	
	(xi)	SEQUENCE I	ESCRIPTION	: SEQ ID NO	: 29:	•	
30	GGCACGAGGA .	AGCTGGAGCG	GGCCGGCGGT	GCAGTCACGG	GGGAGCGAGG	CCTGCTGGGC	60
	TTGGCAACGA	GGGACTCGGC	CTCGGAGGCG	ACCCAGACCA	CACAGAÇACT	GGGTCAAGGA	120
35	GTAAGCAGAG	GATAAACAAC	TGGAAGGAGA	GCAAGCACAA	AGTCATCATG	GCTTCAGCGT	180
33	CTGCTCGTGG	AAACCAAGAT	AAAGATGCCC	ATTTTCCACC	ACCAAGCAAG	CAGAGCCTGT	240
	TGITTTGTCC	AAAATCAAAA	CTGCACATCC	ACAGAGCAGA	GATCTCAAAG	ATTATGCGAG	300
40	AATGTCAGGA	AGAAAGTTTC	TGGAAGAGAG	CTCTGCCTTT	TTCTCTTGTA	AGCATGCTTG	360
	TCACCCAGGG	ACTAGTCTAC	CAAGGTTATT	TGGCAGCTAA	TTCTAGATTT	GGATCATTGC	420
45	CCAAAGTTGC .	ACTTGCTGGT	CTCTTGGGAT	TTGGCCTTGG	AAAGGTATCA	TACATAGGAG	480
	TATGCCAGAG	TAAATTCCAT	TTTTTTGAAG	ATCAGCTCCG	TEEECTECT	TTTGGTCCAC	540
	AGCATAACAG	GCACTGCCTC	CTTACCTGTG	AGGAATGCAA	AATAAAGCAT	GGATTAAGIG	600
50	AGAAGGGAGA	CTCTCAGCCT	TCAGCTTCCT	AAATTCTGTG	TCTGTGACTT	TCGAAGTTTT	660
	TTAAACCTCT	GAATTTGTAC	ACATTTAAAA	TTTCAAGIGT	ACTITAAAAT	AAAATACTTC	720
55	TAATGGAAAA .	ААААААААА	АААААААА	ACTCGA			756

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2100 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

					,		
10	NCCAGAGGCA	GAAAGTCCTG	CTTCTGGGGC	GTAACCTACA	GGATATCCTT	GGAACAGAAG	60
10	ATCTTATTGT	GGAAGTRACT	TCCAATGATG	CTGTGAGATT	TTATCCCTGG	ACCATTGATA	120
	ATAAATACTA	TTCAGCAGAC	ATCAATCTAT	CTCTCCTCCC	AAACAAATTŢ	CTTGTTACTG	180
15	CAGAGATTGC	AGAATCTGTC	CAAGCATTTG	TGGTTTACTT	TGACAGCACA	CAAAAATCGG	240
	GCCTTGATAG	TGTCTCCTCA	TGGCTTCCAC	TGGCAAAAGC	ATGGTTACCY	GAGGTGATGA	300
20	TCTTGGTCTG	CGATAGAGTG	TCTGAAGATG	GTATAAACCG	ACAAAAAGCT	CAAGAATGGT	360
20	GCATCCAAAC	ATGGCTTTGA	ATTGGTAGAA	CTTAGTCCAG	AGGAGTTGCC	TGAGGAGGAT	420
	GATGACTTCC	CAGAATCTAC	AGGAGTAAAG	CGAATTGTCC	AAGCCCTGAA	TGCCAATGTG	480
25	TGGTCCAATG	TAGTGATGAA	GAATGATAGG	AACCAAGGCT	TTAGCTTGCT	GCAACTCATT	540
	GACTGGAACA	AACCATAGCA	TTGGGTCAGC	AGATCCCTGT	CACCCAGAGC	AACCCCATTT	600
30	GCCAGCAGCA	GATAGTACTG	AATCCCTCTC	TGATCATCGG	GGTGGTGCAT	CTAACACAAC	660
	AGATGCCCAG	GTTGATAGCA	TTGTGGATCC	CATGTTAGAT	CTGGATATTC	AAGAATTAGC	720
	CAGTCTTACC	ACTGGAGGAG	GAGATGTGGA	GAATTTTGAA	AGACTCTTTT	CAAAGTTAAA	780
35	GGAAATGAAA	GACAAGGCTG	CGACGCTTCC	TCATGAGCAA	AGAAAAGTGC	ATGCAGAAAA	840
	GGTGGCCAAA	GCATTCTGGA	TGGCAATCGG	GGGAGACAGA	GATGAAATTG	AAGGCCTTTC	900
40	ATCTGATGAA	GAGCACTGAA	TTATTCATAC	TAGGGTTTGA	CCAACAAAGA	TGCTAGCTGT	960
	CTCTGAGATA	CCTCTCTACT	CAGCCCAGTC	ATATTTTGCC	AAAATTGCCC	TTATCATGTT	1020
	GGCTGCCTGA	CTTGTTTATA	GGGTCCCCTT	AATTTTAGTT	TTTAGTAGGA	GGTTAAGGAG	1080
45	AAATCTTTTT	TTTCCTCAGT	ATATTGTAAG	AGAGTGAGGA	ATACAGTGAT	AGTAATGAGT	1140
	GAGGATTTCT	TAAATRTACT	TTTTTTTGT	TCTAGGAATG	AGGGTAGGAT	AAATCTCAGA	1200
50	GETCTGTGTG	ATTTACTCAA	GTTGAAGACA	ACCTCCAGGC	CATTCCTGGT	CAACCTTTTA	1260
	AGTAGCATTT	CCAGCATTCA	CACTTGATAC	TGCACATCAG	GAGTTGTGTC	ACCTITCCTG	1320
	GGTGATTTGG	GTTTTCTCCA	TTCAAGGAGC	TTGTAGCTCT	GAAGCTATGA	TGCTTTTATT	1380
55	GGGAGGAAAG	GAGGCAGCTG	CAGAATIGAT	GTGAGCTATG	TGGGGCCGAA	GTCTCAGCCC	1440
	GCAGCTAAGT	CTCTACCTAA	GAAAATGCCT	CTGGGCATTC	TTTTGAAGTA	TAGTGTCTGA	1500
60	GCTCATGCTA	GAAAGAATCA	AAAAGCCAGT	GTGGATTTTT	AGACTGTAAT	AAATGAGGCA	1560

	107	
	AAGGATTICT ATTCCAGTGG GAAGRAAACC TCTCTACTGA GTTGTGGGGG ATATGTTGTA	1620
	TGTTAGAGAG AACCTTAAGG AGTCCTTGTA TGGGCCATGG AGACAGTATG TGATAACATA	1680
5	CCGTGATTTT CATGAAGAAA TTCTTCTGTC TTAGAGTTCT CCCCTGCTGC TTGAGATGCC	1740
	AGAGCTGTGT TGTTGCACAC CTGCAAAACA AGGCACATTT CCCCCTTTCT CTTTAAAGCC	1800
10	AAAGAGAGAT CACTGCCAAA GTGGGAGCAC TAAGGGGTGG GTGGGGAAGT GAAATGTTAG	1860
10	GCGATGAATT CCTGAGCACC TTGTTTTTCT TCCAAGGTTC GTAGCTCCTC TCTGCCCTTC	1920
	CAAGCCTGTA ACCTCGGAGG ACTATCTTTT GITCTTTATC CITTGTCTTG TTTGAGTGGG	1980
15	TCAGCCCCAG AGGAACTGAT AAGCAAATGG CAAGTTTTTA AAGGAAGAGT GGAAAGTACT	2040
	GCAAATAAAAATCCTTATTT. GTTTTTGTAGAAAAAAAAAAAAAAAAAAA AAAAAAAAAG	2100
20		
	(2) INFORMATION FOR SEQ ID NO: 31:	
25	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1448 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double	
30	(D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:	
	AAAAAAAAA AAAGCCCACC TGAAAGCCTG TCTCTTTCCA CTTTGTTGGC CCTTCCAGTG	
	ARRESTRATES AND CONTROL TO TOTAL TOTAL CONTROL	60

	AAAAAAAAA	AAAGCCCACC	TGAAAGCCTG	TCTCTTTCCA	CITICITICC	CCTTCCAGTG	60
35	GGATTATCGA	GCATGTTGTT	TTTTCATAGT	GCCTTTTTCC	TTATTTCAAG	GGTTGCTTCT	120
	GAGTGGTGTT	TTTTTTTTT	TTAATTTGTT	TTGTTTTAAA	ATAAGTTAAA	GACAGTCCAG	180
	AGCTTTTCAG	CCAATTTGTC	TCCTACTCTG	TGTAAATATT	TTTCCCTCCG	GGCAGGGGAG	240
40	CCAGGGTAGA	GCAAAGGAGA	CAAGCAGGAG	TGGAAGGTGA	GCCTTCTCC	TGCTTGTACT	300
	AAGCCAGGAG	STITAAGCTC	CAGCTTTAAG	GGTTGTGAGC	CCCTTGGGGT	TCAGGGAACT	360
45	GCTTGCCCAG	GGTGCAGTGT	GAGTGTGATG	GCCACCGG	GCAAGAGGGA	AGGTGACCGC	420
	CCAGCTCTCC	CACATCCCAC	TGGATCTGGC	TTACAGGGG	GTCGGAAGCC	TGTCCTCACC	480
	GTCTCGGGGG	TTGTGGCCCC	CGCCCCTCC	CTATATGCAC	CCCTGGAACC	AGCAAGTCCC	540
50	AGACAAGGAG	AGCGGAGGAG	GAAGTCATGG	GAACGCAGCC	TCCAGTTGTA	GCAGGTTTCA	600
	CTATTCCTAT	GCTGGGGTAC	ACAGTGAGAG	TACTCACTTT	TCACTTGTCT	TGCTCTTAGA	660
55	TTGGGCCATG	GCTTTCATCC	TGTGTCCCCT	GACCTGTCCA	GCTCACTCTC	AGGGCAGCAC	720
	TGGGAAGCTG	GAGTGCTGCT	TGTGCCTCCC	TTCCCAGTCG	GCTGTGTTGA	CTGCTGCTCC	780
•	CCACCCCTAC	CGATGGTCCC	AGGAAGCAGG	GAGACTTCCC	GAAGGCAAGA	TTGGAAAGAC	840
60	AGGAAGACCA	AGGCCTCGGC	AGAACTCTCT	GTCTTCTCTC	CACTTCTGGT	CCCCTGTGGT	900

	· ·	
	GATGTGCCTG TAATCTTTTT CTCCACCCAA ACCCCTTCCC ACGACAAAA CAAGACTGCC	960
5	TCCCTCTCTT CCGGGAGCTG GTGACAGCCT TGGGCCTTTC AGTCCCAAAG CGGCCGATGG	1020
J	GAGTCTCCCT CCGACTCCAG ATATGAACAG GGCCCAGGCC TGGAGCGTTT GCTGTGCCAG	1080
	GAGGCGGCAG CTCTTCTGGG CAGAGCCTGT CCCCGCCTTC CCTCACTCTT CCTCATCCTG	1140
10	CTTCTCTTTT CCTCGCAGAT GATAAAAGGA ATCTGGCATT CTACACCTGG ACCATTTGAT	1200
	TGTTTTATTT TGGAATTGGT GTATATCATG AAGCCTTGCT GAACTAAGTT TTGTGTGTAT	1260
15	ATATTTAAAA AAAAAATCAG TGTTTAAATA AAGACCTATG TACTTAATCC TTTAACTCTG	1320
13	COGATAGCAT TTGGTAGGTA GTGATTAACT GTGAATAATA AATACACAAT GAATTCTTMA	1380
	AAAAAAAAA AAAAAAAAA AAAAAAAAA AAACCCCGGG GGGGCCCCCAATT	1440
20	CCCCCCAA	1448
	1	
25	(2) TATEODMATICAL POR CEO TO MO. 22.	
25	(2) INFORMATION FOR SEQ ID NO: 32:	
	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 456 base pairs	
	(B) TYPE: nucleic acid	
30	(C) STRANDEDNESS: double (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:	
35	GGCACAGCAA ACTTGACGCC ATGAAGATCC CGGTCCTTCC TGCCGTGGTG CTCCTCTCCC	60
	TCCTGGTGCT CCACTCTGCC CAGGGAGCCA CCCTGGGTGG TCCTGAGGAA GAAAGCACCA	120
40	TTGAGAATTA TGCGTCACGA CCCGAGGCCT TTAACACCCC GTTCCTGAAC ATCGACAAAT	180
••	TGCGATCTGC GTTTAAGGCT GATGAGTTCC TGAACTGGCA CGCCCTCTTT GAGTCTATCA	240
	AAAGGAAACT TCCTTTCCTC AACTGGGATG CCTTTCCTAA GCTGAAAGGA CTGAGGAGCG	300
45	CAACTCCTGA TGCCCAGTGA CCATGACCTC CACTGGAAGA GGGGGCTAGC GTGAGCGCTG	360
	ATTCTCAACC TACCATAACT CTTTCCTGCC TCAGGAACTC CAATAAAACA TTTTCCATCC	420
50	AAAAAAAAA AAAAAAAAC CCCNGGGGGG GCCCGG	456
<i></i>	•	
	(2) THEOREMATON FOR CEO TO NO. 22.	
55	(2) INFORMATION FOR SEQ ID NO: 33:	
	(i) SPONENCE CHARACTERISTICS:	

(A) LENGTH: 1326 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

5	GGCACGAGTG	CAGGCCCAGA	GAGGACTCAT	TGAAAGGACT	GAAAGGGGAG	GTGGCGTTTT	60
,	CTTCCTACCC	AAACTTACCC	CTGTGAGCTG	GACAGCTTGG	TAGCACCTGC	CTGGACTTAG	120
	ATGGTGGTAG	CCAAGAAGAC	TGACATTITA	GGGAACAGGA	CGGGGAGGAG	AAGGCTCTGG	180
10	CACACACACA	TGTGTCCATA	TGTCCTGCAA	TGGTCTGGGG	ACTATTGCTA	GGCTAGGAGC	240
	CCTAAGTGTC	TTCTTCCTCA	TGTCTMTTCT	CCCCTGTSTC	ATGGGCCCTA	AGRICICITY	300
15	CACTGGGCCT	GCCTCAATGA	ACGTGCTGCC	CAGCTACCCC	GAAACACGGC	ANCTGCCGGC	360
15	TATCAATGCC	CCAGCTGCAA	TGGCCCATCT	TCCCCCAACC	AACCTGGCTG	GCCCGTGGG	420
	CTCCGCACTG	AGARARAAAS	TTGGCACART	CAACTGGGCC	CGGGCAGGAC	TGGGCCYCCC	480
20	TCTGATCGAT	GAAGKTGGTG	ARCCCAGAGC	CCGAGCCCCT	CAACACGTCT	GACTTCTCTG	540
	ACTGGTCTAG	TTTTAATGCC	AGCAGTACCC	CTGGACCAGA	GGAGGTAGAC	AGCGCCTCTG	600
25	CTGCCCCAGC	CTTCTACAGC	CGAGCCCCCC	GCCCCCAGC	TTCCCCAGGC	CGGCCCGAGC	660
23	AGCACACAGT	GATCCACATG	GGCAATCCTG	AGCCCTTGAC	TCACGCCCCT	AGGAAGGTGT	720
	ATGATACGCG	GGATGATGAC	CGGACACCAG	GCCTCCATGG	AGACTGTGAC	GATGACAAGT	780
30	ACCGACGTCG	GCCGGCCTTG	GCTTGCCTGG	CCCGGCTGCT	AAGGAGCCGG	GCTGGGTCTC	840
	GGAAGCGRCC	GCTGACCCTG	CTCCAGCGGG	CGGGGCTGCT	GCTACTCTTG	GGACTGCTGG	900
35	GCTTCCTGGC	CCTCCTTCCC	CTCATGTCTC	GCCTAGGCCG	GCCGCAGCT	GACAGCGATC	960
,,	CCAACCTGGA	CCCACTCATG	AACCCTCACA	TCCGCGTGGG	CCCCTCCTGA	GCCCCCTTGC	1020
	TTGTGGCTAG	GCCAGCCTAG	GATGTGGGTT	CTGTGGAGGA	GAGGCGGGGT	AATGGGGAGG	1080
40	CTGAGGGCAC	CTCTTCACTG	CCCCTCTCCC	TCAAGCCTAA	GACACTAAGA	CCCCAGACCC	1140
	AAAGCCAAGT	CCACCAGAGT	GGCTGCAGGC	CAGGCCTGGA	GTCCCCGTGG	GTCAAGCATT	1200
15	TGTCTTGACT	TGCTTTCCTC	CCGGGTYTCC	AGCCTCCGAC	CCCTCGCCCC	ATGAAGGAGC	1260
	TGGCAGGTGG	AAATAAACAA	CAACTTTATT	AAAAAAAA	АААААААА	АААААААА	1320
	AAANAA						1326

50

(2) INFORMATION FOR SEQ ID NO: 34:

55 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 710 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

660

	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:	
	GCGAAAGAGA AAAAGGCTGG AGCTCCCGCC CCCGGGGCTG TCAGATGGCT TGGGFFTCTG	60
5	CGACGCGATT GGCTCGCGGA GGGCAGAAAT TACTCAGCAA ACATGACTAT TATTAGCTGC	120
	TTAGCAACAG CTCACCAAAG TAGAGAGACC ACCCAGGTAG GCAACCCAGT GTGTGCATCC	180
10	TCGGCTTCGG GGCAGCCTCT GAGAGCGCCA ACCTTCTCGC ATGCAATACT TCCATTAAGG	240
10	AATGCTCCCC CTCCTTTCTC TCTTATTCCT TTTCTTTTCA ACAGTGTCTT CTTTTTGTGG	300
	GATGCCTTTG CGCGCACACA CGCGCGCGCA SGCACACACA CGAACATTTG CCTCGCGGTA	360
15	GACACGGGG GAAATGTWAT ATTTTTTAA GCGCTTAAAC AATTTCTGAA ATTCCTCAAA	420
_	GAAAAGCCTT TCAGARGCAC CTTGGCCTCA AGCTGCAACA AATACTGGGA RGTCCGGCTC	480
20	GCATTCCCAG GCCTGCACCA ATAATGACAG CGTGCTGGAT ARTGCGCCAG TGTGTGCCAG	540
20	ATTITITIT CCTCTTCTCT TTTCTTTTAT AACTAAAGGG AAGACTTAGG CTCTTGCAGG	600
	GAACAACGCC TCGCATTAAG ATAAACAGAA TGGAAAGTTA AAGAGGAAAG CAAGGACGTT	660
25	GGGAAAAGCC ATCTTTCTTA AAATCCGTCT GCCCCCCAGC CGCTTTCTCC	710
30	(2) INFORMATION FOR SEQ ID NO: 35:	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 1188 base pairs (B) TYPE: nucleic acid	
35	(C) STRANDEDNESS: double (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:	
40	GATGGCTTTT ATATCTATTA TCGACCCACA GACAGTGACA ATGATAGTGA CTACAAGAAG	60
	GATATGGTGG AAGGGGACAA GTACTGGCAC TCCATCAGCC ACCTGCAGCC AGAGACCTCC	120
A.E.	TACGACATTA AGATGCAGTG CTTCAATGAA GGAGGGGAGA GCGAGTTCAG CAACGTGATG	180
45	ATCTGTGAGA CCAAAGCTCG GAAGTCTTCT GGCCAGCCTG GTCGACTGCC ACCCCCAACT	240
	CTGGCCCCAC CACAGCCGCC CCTTCCTGAA ACCATAGAGC GGCCGGTGGG CACTGGGGCC	300
50	ATGGTGGCTC GCTCCAGCGA CCTGCCCTAT CTGATTGTCG GGGTCGTCCT GGGCTCCATC	360
	GTTCTCATCA TCGTCACCTT CATCCCCTTC TGCTTGTGGA GGGCCTGGTC TAAGCAAAAA	420
55	CATACAACAG ACCTGGGTTT TCCTCGAAGT GCCCTTCCAC CCTCCTGCCC GTATACTATG	480
. 1. 1		

GTGGACGGC CTGTGCTAAT GGGATCCACA TGAATAGGGG CTGCCCCTCG GCTGCAGTGG

GCTACCCGGG CATGAAGCCC CAGCAGCACT GCCCAGGCGA GCTTCAGCAG CAGAGTGACA

	CCAGCAGCCT GCTGAGGCAG ACCCATCTTG GCAATGGATA TGACCCCCAA AGTCACCAGA	720
5	TCACGAGGG TCCCAAGTCT AGCCCGGACG AGGGCTCTTT CTTATACACA CTGCCCGACG	780
	ACTCCACTCA CCAGCTGCTG CAGCCCCATC ACGACTGCTG CCAACGCCAG GAGCAGCCTG	840
	CTGSTGTGGG CCAGTCAGGG GTGAGGAGGAG CCCCCGACAG TCCTGTCCTG	900
10	GGGACCCTCC ATTTCACTCA GGGCCCCCAT GCTGCTTGGG CCTTGTGCCA GTTGAAGAGG	960
	TOGACAGTCC TGACTCCTGC CAAGTGAGTG GAGGAGACTG GTGTCCCCAG CACCCCGTAG	1020
15	GGGCCTACGT AGGACAGGAA CCTGGAATGC AGCTCTCCCC GGGGCCACTG GTGCGTGTGT	1080
	CTTTTGAAAC ACCACCTCTC ACAATTTAGG CAGAAGCTGA TATCCCAGAA AGACTATATA	1140
	TTGTTTTTT TTTAAAAAAA AAAAAAAAA AWCYCGGGG GGGCCCC	1188
20		
	(2) INFORMATION FOR SEQ ID NO: 36:	
25	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 956 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double	
30	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:	
	GGCAGAGCAG TGAAAATGCA TCCTAAAAAT TCAATGTTTA TACCAGGCTC ATGACACTAA	60
35	GATGTGACAT CTGGACACGA GGGGTCAGCC ACGTGGATAC ATCCCTCCCA GATTGCATCT	120
	CCAGGAATCA CTCTGCTAGC AGAATGGGCG CCCCATCCCT TACTATGCTG CTCCTCCA	180
40	AAGTGCAGCC CAGAAGGACC CAGGCCTTTG ATGCACATTG GGTGGGTCTC CCACTACTTT	240
	AGTTGAAATG GGAGCATGCT GGAGTCGGCG TTCTGTTGCT TCTGGTGAGA AGGACATCCC	300
	ATTGACCCCT GGCCACCAGG TCCAGTATTC CATCCTTCCT TCTGTCCCAG CCTATCGCCC	360
45	TCCCCACYAG GCCCACCCCC ACAACTTCTC CTCAAGGGAG GTTNTCCCGC AGCTGGAGGG	420
	CTTGCACAGA CCAGCAGTCA CAGAAATCAT TCTTCCTGCT GTACTGGGCC TTAACTGCCT	480
50	GCAAATGTCC GAGCACTACT GCATAGGATG CCAGAGCCAC CGAAGATAAA CACAGCCAAG	540
	TTTAATAATA ATAAAAGGAA AAATCTCAGC CTGCAGAACT CTGGTTTTGA CCCACCATCG	600
	GCCAGATGCA CATCTTCAGG GCCTGTTGAG CACCTTCTGA AAAGCAGGGC TCGTAATAGA	660
55	CTCCAGCACA TTCCATCAGA GTCAGGAAAA CTGCGGTGAG TCCCAGAGAA TCTAGGGTGC	720
	AGGGCAGGGA GCAGGAGTCA TAAGGAGTGA TAACCTAAAC TGTGTGTAGT CAGCGGGGGAG	780
60	GGTCTTATGT TATCAGGTGA AATGAGAGCC AGTAAGTTAG TTGATCCTGT CACAGATATA	840

	indication discontinuous announced controlled detection in incident controlled detection in inciden	900
	ACATGGTTCT TCTGTTCCAC AGACATTAAA GGGGCTTTCT GCAATTACTT AAAAAA	956
5		
	(2) INFORMATION FOR SEQ ID NO: 37:	
10	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1603 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double (D) TOPOLOGY: linear	
15	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:	
20	TCGACCCACG CGTCCGCTCT GCCAGGAATC TGGTCTTTCT GTAGACCCAA GTCAGAAAGA	60
20	ACCATTIGIG GAGITAAATC GAATATTAGA RGCATTAAAR GICAGAGITC IGAGACCIGC	120
	TCTGGAATGG GCAGTTTCAA ACCGAGAGAT GCTTATAGCC CAAAACAGCT CCTTGGAATT	180
25	TAAACTACAC AGACTGTATT TTATTAGCTT RITAATGGGT GGAACACAAA TCAGCGAGAR	240
	GCATTACAAT ATGCTAAAAA TTTTCAGCCA TTTGCCCTAA ATCATCAAAA AGACATTCAG	300
	GITTIGATGG GAAGCCTIGI GTACCTGAGA CAAGGGATTG AGAACTCACC ATATGTTCAC	360
30	CTACTTGATG CAAACCAGTG GGCTGATATC TGTGACATCT TTACACGGGA TGCTTGTGCC	420
	CTCCTGGGGC TCTCCGTGGA GTCCCCTCTC AGTGTCAGTI TCTCAGCAGG TTGTGTGGCG	480
35	CTGCCAGCTT TAATTAACAT CAAAGCCGTG ATTGAACAGA GGCAGTGTAC TGGAGTTTGG	, 540
33	AACCAGAAAG ATGAATTACC TATTGAAGTG GACCTTGGTA AAAAGTGCTG GTATCACTCT	600
	ATATTTGCCT GCCCCATTCT TCGTCAGCAA ACAACAGATA ACAATCCACC CATGAAATTG	660
40	GTCTGTGGTC ATATTATATC AAGAGATGCC CTGAATAAAA TGTTTAATGG TAGCAAATTA	720
	AAATGTCCCT ACTGTCCAAT GGAACAAAGT CCAGGAGATG CCAAACAGAT ATTTTTCTGA	780
45	AGAGATAACT TTAGTTTGCA ATTTGTAAGT GAAACTGAAT CGTGGGTGCA TTTCAGAAGA	840
45	GAACGITCCA TATAATGCAG CTAACCAAGG ACTCCTGTGT TTCTATAAGC TAATGCTCCA	900
•	GAAACTITGC CAACCTGITA GTGTACACAC ACTGAGGGGA GTGCTCCCGG TGAATATTAT	960
50	CATAGGGCIT TATTATATIC TIGGTCTICA TITCTGATCA AGTAAATACA CCAGCAGTIG	1020
	TCATTCAATG CAGGITTITG TACTTAATTA TATGGTGATT TTTTTACTIT TTAAGAGCAG	1080
	AAACGGAAAT TGACCTCCCC GCCATGTGTT TAATATTCCT CCTGCTTTTA CTTTTGTCAT	1140
55	TITCITGATA ATCGTAAGCC TIGAGAGIGT TIGIGAAAAA GITITATTIC CIGITATGTA	1200
	TACATAATTA AATGAAAATT CTTCAGAAAA AGTTTGATAA ATTGAATTGT GGTTATGAAA	
60	CTAATTIGCA TITITATITG CITAAGAAAG AAAGCTGIGA TAGATTCCAG ATATGCTTTT	
	THE TENED THE PROPERTY OF THE	1340

	TGATGTTTTC	CTCTGCTCCA	GCTCCAAGAA	GTCAGCACAC	CTGCATTTTA	GCTCTGCATG	1380
5	CAGCCCCAGC	AGGCTGCGTG	TTTAAGAATT	TCATTGTTTA	ACTGGCTGGT	GTGAĠAAGTC	1440
5	TTCCGTTAGC	ATAGAGTGGA	AGGAGTACTA	TIGITIGGIT	GGGTTTTTGT	TIGITIGITT	1500
	TTTGTTTTTG	CTTTTATTGC	CAAGAGGTGC	TTGTTTTAAA	AGTATGTTTA	ATAAAATGAA	1560
10	ATTCTAAAGT	TAARAAGTGT	TCTTAAAGTT	GATATTTAAC	TCT		1603

15 (2) INFORMATION FOR SEQ ID NO: 38:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1089 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

				•			
25	GGCACGAGCT	ACCTTTCTGC	CIGCTITICCT	GGCTGCAACA	GCACGAATCT	CACGGGCTGT	60
	GCGTGCCTCA	CCACCGTCCC	TGCTGAGAAC	GCAACCGTGG	TTCCTGGAAA	ATGCCCCAGT	120
30	CCTGGGTGCC	AAGAGGCCTT	CCTCACTITC	CTCTGTGTGA	TGTGTATCTG	CAGCCTGATC	180
50	GGTGCCATGG	CAAGACACCC	TCAGTCATCA	TCCTCATCAG	GACAGTCAGC	CCTGAACTCA	240
	AGTCTTACGC	TTTGGGAGTT	CTTTTTCTCC	TCCTTCGTTT	GTTGGGCTTC	ATCCCTCCAC	, 300
35	CCCTCATCTT	CGGGCTGGC	ATCGACTCCA	CCTGCCTGTT	CTGGAGCACG	TTCTGTGGGG	360
	AGCAAGGCGC	CTGCGTCCTC	TACGACAATG	TGGTCTACCG	ATACCTGTAT	GTCAGCATCG	420
40	CCATCGCGCT	CAAATCCTTC	GCCTTCATCC	TGTACACCAC	CACGTGGCAG	TGCTGAGGAA	480
40	AAACTATAAA	CGCTACATCA	AAAACCACGA	GGCGGGCTG	AGCACCAGTG	AGTTCTTTGC	540
	CTCTACTCTG	ACCCTAGACA	ACCTGGGGAG	GGACCCTGTG	CCCGCAAACC	AGACACATAG	600
45	GACAAAGTTT	ATCTATAACC	TGGAAGACCA	TGAGTGGTGT	GAAAACATGG	AGTCCGTTTT	660
	ATAGTGACTA	AAGGAGGCT	GAACTCTGTA	TTAGTAATCC	AAGGGTCATT	ТТТТСТТАА	720
50	AAAAGAAAA	AAAGGTTCCA	AAAAAAACCA	AAACTCAGTA	CACACACACA	GGCACAGATG	780
50	CACACACACG	CAGACAGACA	CACCGACTTT	GTCCTTTTTC	TCAGCATCAG	AGCCAGACAG	840
	GATTCAGAAT	AAGGAGAGAA	TGACATCGTG	CGGCAGGGTC	CTGGAGGCCA	CTCGCGCGGC	900
55	TGGGCCACAG	AGTCTACTTT	GAAGGCACCT	CATGGTTTTC	AGGATGCTGA	CAGCTGCAAG	960
	CAACAGGCAC	TGCCAAATTC	AGGGAACAGT	GGTGGCCAGC	TTGGAGGATG	GACATTTCTG	1020
60	GATACACATA	CACATACAAA	ACAGAAAACA	TTTTTTAAAA	GAAGTTTCCT	ААААТАААА	1080
UU							

	АААААААА	÷		1089
٠ -		, ,		
5	(2) INFORMATION FOR SEQ ID NO: 39:			
10	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 629 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	1	1 .	
15	(xi) SEQUENCE DESCRIPTION: SEQ ID NO	: 39:		
13	AGCTCAGTTC CCTTAGAAAT GAAATTTTAA ATGACACTAC	CAGGTAAGCC	ACTGAGACCA	60
	GTGGAGGTGA TAGCTAAGAA CATAAGGAAT TAAGAATTTT	TAATGGAGAA	AGGAGGTAAT	120
20	GAATACCAGT TACATCCTAA GACTCACTGT AGTGGTGAGT	GTTGTAATTT	ATCTCGCTAT	180
	CCATCCTCTT TTAAGTTTTT CCTTAGAAAG TCCTCTATTG	GTACCTTGGA	GGGACTCCTG	240
25	TCAAAATATA TGGAAAAGTG GGTCTGTGTG GTACAAGAGG	TGGACTTTGC	CACACATGGA	300
	AGFFTGCTGC CAAGATCTTC ACTAATGAAA GAAATCACCA	GTGAGCTGCA	CAGATTAGCC	360
	AAATACTGAG CTCATTAGAA CTACTAAGGC CTGGACATTT	CTGCCTAATC	CAGGACTCCT	420
30	GIAATTATCA GICTITGCTT TGGAGCTTCC CATTGTGTAG	CTGARAATTT	GTCATATCTG	480
	CATTATAATC TAAGGCTCCA CATACTTAAT CCTGCTTCTC	CCCCTTTTTC	TTTCCCTTTC	540
35	CCAGCGGTCA GCTCTGCTGC ATAGTCTGAA GACTTTCCCT	GCCCAATCCT	GATAAAATTC	600
	TIGCACTOGT AACCCCATCT CAGTGTCTG			629
40	(2) INFORMATION FOR SEQ ID NO: 40:			•
45	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1964 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear			
50	(xi) SEQUENCE DESCRIPTION: SEQ ID NO	: 40:		
J U	AAGAAGACAT GGAAATTGCT GAAGGATGTT TCAGGCATAT	TAAGAAAATC	TTTACGCAGC	60
	TTGAGGAATT CAGAGCCTCT GAATTGCTTC GAAGTGGACT	GGACAGATCT	AAATACCTTT	120
55	TAGTGAAAGA AGCCAAAATT ATTGCTATGA CCTGTACTCA	TGCTGCCTTA	AAACGACATG	180
	ACTIGGICAA GCTAGGITIC AAGTATGACA ACATTITGAT	GGAAGAGGCT	GCTCAGATTC	240
60	TOGAGATAGA AACTITITATC CCTCTTCTTC TACAGAATCC	TCAGGATGGA	TTTAGCCGAC	300

	TAAAACGATG	GATTATGATT	GGCGATCATC	ACCAGITACC	TCCAGTTATT	AANGAACATG	360
	GCCTTTCAAA	AGTACTCAAA	CATGGAGCAG	TCTCTCTTCA	CTCGCTTTGT	TCGCGTTGGA	420
5	GTTCCGACTG	TTGACCTTGA	TGCTCAAGGG	AGAGCCAGAG	CAAGCTTGTG	CAMCTNCTAC	480
	AACTGGCGAT	ACAAGAATCT	AGGAAACTTA	CCCCATGTGC	AGCTCTTGCC	AGAGTTTAGT	540
10	ACAGCAAATG	CTGGCTTACT	GTATGACTTC	CAGCTCATTA	ATGITGAAGA	TTTTCAAGGA	600
•	GTGGGAGAAT	CTGAACCTAA	TCCTTACTTC	TATCAGAATC	TTGGAGAGGC	AGAATATGTA	660
	GTAGCACTTT	TTATGTACAT	GTGTTTACTT	GGTTACCCTG	CTGACAAAAT	CAGTATTCTA	720
15	ACAACATATA	ATGGCCAAAA	GCATCTTATT	CGCGACATCA	TCAATAGACG	ATGTGGAAAC	780
	AATCCATTGA	TTGGAAGACC	AAACAAGGTG	ACAACTGTTG	ATAGATTICA	AGGTCAACAG	840
20	AATGACTATA	TTCTTCTTTC	TCTCGTACGA	ACCAGGGCAG	TGGGCCATCT	GAGGGATGTC	900
	CGTCGCTTGG	TAGTGGCCAT	GTCTAGAGCC	AGACTTGGAC	TTTATATCTT	CGCCAGAGTA	960
	TCCCTCTTCC	AAAACTGTTT	TGAACTGACT	CCAGCTTTCA	GTCAGCTCAC	AGCTCGCCCC	1020
25	CTTCATTTGC	ATATAATTCC	AACAGAACCT	TTCCCAACTA	CTAGAAAGAA	TGGAGAGAGA	1080
	CCATCTCATG	AAGTACAAAT	AATAAAAAT	ATGCCCCAGA	TGGCAAACTT	TGTATACAAC	1140
30	ATGTACATGC	ATTIGATACA	GACTACACAT	CATTATCATC	AGACTTTATT	ACAACTACCA	1200
_	CCTGCTATGG	TAGAAGAGGG	TGAGGAAGTT	CAAAATCAAG	AAACAGAATT	GGAAACAGAA	1260
	GAAGAGGCCA	TGACTGTTCA	AGCTGACATC	ATACCCAGTC	CAACAGÁCAC	CAGCTGCCGT	1320
35	CAAGAAACTC	CAGCCTTTCA	AACTGACACC	ACCCCCAGTG	AGACAGGAGC	CACTTCCACT	1380
	CCAGAAGCCA	TCCCTGCTTT	ATCTGAGACC	ACCCCTACTG	TGGTAGGAGC	TGTATCTGCA	1440
40	CCGGCAGAAG	CTAACACACC	TCAGGATGCC	ACATCTGCCC	CAGAAGAGAC	CAAGTAGCCA	1500
	AACTGTAGTC	CTTCTAAAGG	AGGACATGGC	AGTCAAAAAG	TCTGAGTAAA	CCTGTTTTTT	1560
	GTATTTTATA	TTTGCTTCTG	CCATTTTACT	GTCACTAATT	AATGTTTAGT	TCTTATATTT	1620
45	GTTAACTGAT	TICGGIGICT	TGAATATATT	TTTTTAAATT	ATGTGTATGA	ACAATTCTAG	1680
	TITCATTIGT	TCAATCAGAA	GAGCAAATAA	CCATTCCTTT	CATGITTIGA	TCACTGAGTG	1740
50	TGTCTGTAAT	CATACCTACA	TTAAAATCAT	TTTCTATGAA	TATATAATAT	ATACTTCACA	1800
	TTTTTAGTGA	ACTTCTCTAA	AGAAGAGGAC	AGAATATACT	GGACTTAACC	ACGAATACCC	1860
	TTGAGTGTCC	AAATTGGGAA	GGAACTKGTT	TCTTCYGTTA	TACTAYCAAA	TGCTTAAATT	1920
55	CKGTTTCCTT	TTTTCTTACC	TTTGTTTGCT	GICTTIATGT	AAAG		1964

60 (2) INFORMATION FOR SEQ ID NO: 41:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1522 base pairs

.(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ'ID NO: 41:

10	CGTGTCCGCG	CGCCTGGGAG	ACCCTGCCTC	GCCCCGGACG	ceccececc	CCCGCGGCTG	60
	GAGGGTGGTC	GCCACTGGGA	CACTGTGAAC	CAGGAGTRAG	TCGGAGCTGC	CCCCTCCCC	120
15	AGGCCATGGA	CTGTGAGGTC	AACAACGGTT	CCAGCCTCAG	GGATGAGTGC	ATCACAAACC	180
13	TACTGGTGTT	TGGCTTCCTC	CAAAGCTGTT	CTGACAACAG	CTTCCGCAGA	GAGCTGGACG	240
	CACTGGGCCA	CGAGCTGCCA	GTGCTGGCTC	CCCAGTGGGA	GGGCTACGAT	GAGCTGCAGA	300
20	CTGATGGCAA	CCGCAGCAGC	CACTCCCGCT	TGGGAAGAAT	AGAGGCAGAT	TCTGAAAGTC	360
	AAGAAGACAT	CATCCGGAAT	ATTGCCAGGC	ACCTCGCCCA	GGTCGGGGAC	AGCATGGACC	420
25	GTAGCATCCC	TCCGGGCCTG	GTGAACGGCC	TGGCCCTGCA	GCTCAGGAAC	ACCAGCCGGT	480
23	CGGAGGAGGA	CCGGAACAGG	GACCTGGCCA	CTGCCCTGGA	GCAGCTGCTG	CAGGCCTACC	540,
	CTAGAGACAT	GGAGAAGGAG	AAGACCATGC	TGGTGCTGGC	CCTGCTGCTG	GCCAAGAAGG	600
30	TGGCCAGTCA	CACGCCGTCC	TTGCTCCGTG	ATGTCTTTCA	CACAACAGTG	AATTTTATTA	660
	ACCAGAACCT	ACGCACCTAC	GTGAGGAGCT	TAGCCAGAAA	TGGGATGGAC	TGAACGGACA	720
35	GTTCCAGAAG	TGTGACTGGC	TAAAGCTCGA	TGTGGTCACA	GCTGTATAGC	TGCTTCCAGT	780
33	GTAGACGGAG	CCCTGGCATG	TCAACAGCGT	TCCTAGAGAA	GACAGGCTGG	AAGATAGCTG	840
	TGACTTCTAT	TTTAAAGACA	ATGITAAACT	TATAACCCAC	TTTAAAATAT	CTACATTAAT	900
40	ATACTTGAAT	GAAAATGTCC	ATTTACACGT	ATTTGAATGG	CCTTCATATC	ATCCACACAT	960
	GAATCTGCAC	ATCTGTAAAT	CTACACACGG	TGCCTTTATT	TCCACTGTGC	AGGTTCCCAC	1020
45	TTAAAAATTA	AATTGGAAAG	CAGGTTTCAA	GGAAGTAGAA	ACAAAATACA	ATTTTTTTGG	1080
43	ТАААААААА	TTACTGTTTA	TTAAAGTACA	ACCATAGAGG	ATGGTCTTAC	AGCAGGCAGT	1140
	ATCCTGTTTG	AGGAAAGCAA	GAATCAGAGA	AGGAACATAC	CCCTTACAAA	TGAAAAATTC	1200
50	CACTCAAAAT	AGGGACTATC	YATCTTAATA	CTAAGGAACC	AACAATCTTC	CTGTTTAAAA	1260
	AACCACATGG	CACAGAGATT	CNGAACTAAA	GTGCTGCACT	CAAATGATGG	GAAGTCCCGG	1320
55	CCCCAGTACA	CCAGGGGCTT	TGGACTTTTT	TCAACTTCGT	TICCITIIGI	TIGGANTCCA	1380
JJ	AAAGAACCAC	TTTGTGGTTC	TTAAAAGGGT	GTGAAGGTGA	TTTAAGGGGC	CCAGGTCAGC	1440
	CACTGGTTGG	TTTACAAAAT	CNGGGTAACT	AACTGCATAC	AACTITITICC	CNITICCATG	1500
60	NCATCAGGAC	TITGCTAAAG	AC				1522

5	(2) INFORMATION FOR SEQ ID NO: 42:	
10	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 875 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:	
15	TGGGATTICC CITTATCATG GAGGCCITGT CCCACTTCCT CTATGTCCCT TTCCTTGGTG	60
	TCTGTGTCTG TGGGGCCATC TACACTGGCC TGTTCCTTCC TGAGACCAAA GGCAAGACCT	120
20	TCCAAGAGAT CTCCGAGGAA TTACACAGAC TCAACTTCCC CAGGCGGGCC CAGGGCCCCA	180
20	CGTGGAGGAG CCTGGAGGTT ATCCAGTCAA CAGAACTCTA GTCCCAAAGG GGTGGCCGTA	240
	GCCAAAGCCA GCTACCGTCC TGTCCTCTGC TTCCTGCCAG GGCCCTGGTC CTCAMTYCCT	300
25	YCTGCATTCC TCATTTAAGG AGTGTTTATT GAGCACCCTT TGTGTGCAGA CATGGCTCCA	360
	GGTGCTTAGC AATCAWIGGT GAGCGTGGTA TCCAGGCTAA AGGTAATTAA CTGACAGRAA	420
30	ATCAGTAACA ACATAATTAC AGGYTGGTTG TGGCAGYTCA TGACTGTAAT CCCAGCACTT	480
30	TTGGGAGCCA AGGTGGGARG ATCAATTGAG GCCAGAGTTT GAAAMCAGCT AGGTAACATA	540
	GTGAGACCCC CTATCTCTAC AAAAAATTTT AAACATTAGC TGGGCATGGT GGTATGTGCT	600
35	AACAGCTCTA GCTACTCAGG AGGCTGAGGC AGCAGGATCA CTTGAGTCCA AGAGTTCAAG	660
	GTAGCAGTAA GCTACAATCA CACCACTGCA TGCCAGACTG GGTGACAGAG GGAGACTTCA	720
40	TCTCTTTAAA ACATAATAAT AATAATTACA GACTCAGGAA ATGCAGTGAA AGAAAAATAC	780
40	AGGITGGCCA GGTGAGGTGG CTGATGCCTG TAATCCCAGC ACTTTGGGAG GCCAAGATGG	840
	GAAGATTGCT TTGAGACCAG AAGTTTGAGA CCAGC	875
45		
	(2) INFORMATION FOR SEQ ID NO: 43:	
50	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 843 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
55	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:	
	(X1) SEQUENCE DESCRIPTION: SEQ ID NO: 43: CCCACGCGGT CCGNATCGTC CTTCCCTCAC TTCAGAGGGT GGCCAGAGCT GAATACCCAG	60
60	ACACCACAA CTAACCCTCC ACTTCCAAAA CATCATGAGG ATGTATCATC CCACGAGCT	120

	CACCTGACAG	TTACAGAGGA	AACCCGCACC	CAGAATGCAC	GIGCTGICIT	ATGGGAACAC	180
5	TCAGCGCAGA	GTGCTCAGGT	CCGGCCACAC	TCGGGCTGTG	CTTGGTCGTG	CCATGGAATT	240
	CCTCAGGACT	TTCTCAGCCT	CCCTAATGGC	AGAAGCCCCT	TTACAGCAAG	ACATTTACCG	300
	TTTGTCTGAA	AATAGCCGAA	CTGAGCTTTT	CTTCAGGCTA	TATGAGAAGT	CTCTAGACAG	360
10	TGGGCACCGT	CAGAAAGCCC	AGAGCCTTGT	GATAGCTCCC	ACCCTGCCTG	GCTCAGATCT	420
	TCCCATTTTT	TTTCCTCTGG	CACTAACCTC	ACCTTTTGTT	TTTTTGTGTT	TGTGTTTGTT	480
15	TTTGTTTTTG	CAGAGTTGGA	TTACAGAAAC	TCCTATGAAA	TTGAATATAT	GGAGAAAATT	540
	GCCTCCTCCT	TACCTGTAAG	TTCGTCTGCC	TCGGGCCACT	TAGGGGACTC	GCTTTCCTGC	600
	CTTCAGGGGC	CTCCTCCCCT	GTGCAGAGTG	TCTCTGGGAG	CTCAGACCCC	AAATCGAGTG	660
20	TTTTCTGTGT	ACACAGCTTC	CCGGGTGCAC	AGCAATGATG	GACTGGGGCT	GGGGGTTGA	720
	GGTTTGTACT	CAATCCACTT	CGTTTGACAT	TTTCAGGGAG	AAAATGATAG	AATACAATTA	780
25	GACGTCCTGC	AGAATTACTT	TCCTAGACTG	AGAAAGAGCT	AGAGATTTCT	AAAAAAATT	840
	AAA	•	1				843
30	(2) INFORMA	ATION FOR SE	Q ID NO: 44	:			
35	(i)	(B) TYPE (C) STRA	ARACTERISTI STH: 489 bas S: nucleic a ANDEDNESS: 6 DLOGY: line	se pairs acid double			
40	(xi)	SEQUENCE D	ESCRIPTION:	SEQ ID NO:	44:		
	Cattering Coccur	MMC33CC3Mm					

CTCTTAGGCT TTGAAGCATT TTTGTCTGTG CTCCCTGATC TTCAGGTCAC CACCATGAAG 60 TTCTTAGCAG TCCTGGTACT CTTGGGAGTT TCCATCTTTC TGGTCTCTGC CCAGAATCCG 120 45 ACAACAGCTG CTCCAGCTGA CACGTATCCA GCTACTGGTC CTGCTGATGA TGAAGCCCCT 180 GATGCTGAAA CCACTGCTGC TGCAACCACT GCGACCACTG CTGCTCCTAC CACTGCAACC 240 ACCGCTGCTT CTACCACTGC TCGTAAAGAC ATTCCAGTTT TACCCAAATG GGTTGGGGAT 300 50 CTCCCGAATG GTAGAGTGTG TCCCTGAGAT GGAATCAGCT TGAGTCTTCT GCAATTGGTC 360 ACAACTATTC ATGCTTCCTG TGATTTCATC CAACTACTTA CCTTGCCTAC GATATCCCCT 420 55 480 АААААААА 489

	(2) INFORMATION FOR SEQ ID NO: 45:	
5	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 534 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
10	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:	
	GAAGCAGTGT GTATCTATGA TTATATCTCT GTTCATCTAT ATATTTTTGA CATGTAGCAA	60
15	CACCTCTCCA TCTTATCAAG GAACTCAACT CGGTCTGGGT CTCCCCAGTG CCCAGTGGTG	120
13	GCCTTTGACA GGTAGGAGGA TGCAGTGCTG CAGGCTATTT TGTTTTTTGT TACAAAACTG	180
	TCTTTTCCCT TTTCCCCTCC ACCTGATTCA GCATGATCCC TGTGAGCTGG TTCTCACAAT	240
20	CTCCTGGGAC TGGGCTGAGG CAGGGGCTTC GCTCTATTCT CCCTAACCAT ACTGTCTTCC	300
	TITICCCCTTG CCACTTAGCA GITATCCCCC CAGCTATGCC TTCTCCCTCC CTCCCTTGCC	360
25	CTGGCATATA TTGTGCCTTA TTTATGCTGC AAATATAACA TTAAACTATC AAGTGAAAAA	420
2 3	AAAAAAAA AAAACTCCAA GGGGGGCCG GTACCCAATT CCCCCTATAN TGAGTCNTAT	480
	TACAATTCAC TGGGCCGTCG TTTTACAACG TCGTGAATGG GAAAACCTGG GCGT	534
30		
	(2) INFORMATION FOR SEQ ID NO: 46:	
35	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1374 base pairs	
	(B) TYPE: nucleic acid (C) STRANDEDNESS: double	
40	(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 46:	
	GGCACGAGTC CGGGATGAGC TCAGCCGCGG CCGACCACTG GGCGTGGTTG CTGGTGCTCA	6
45	GCTTCGTGTT TGGATGCAAT GTTCTTAGGA TCCTCCTCCC GTCCTTCTCA TCCTTCATGT	12
	CCAGGGTGCT GCAGAAGGAC GCGGAGCAGG AGTCACAGAT GAGAGCGGAG ATCCAGGACA	18
50	TGAAGCAGGA GCTCTCCACA GTCAACATGA TGGACGAGTT TGCCAGATAT GCCAGGCTGG	24
	AAAGAAAGAT CAACAAGATG ACGGATAAGC TCAAAAACCCA TGTGAAAGCT CGGACAGCTC	30
	AATTAGCCAA GATAAAATGG GTGATAAGTG TCGCTTTCTA CGTATTGCAG GCTGCCCTGA	36
55	TGATCTCACT CATTTGGAAG TATTATTCTG TCCCTGTGGC TGTCGTGCCG AGTAAATGGA	42
	TAACCCCTCT AGACCGCCTG GTAGCCTTTC CTACTAGAGT AGCAGGTGGT GTTGGAATTA	48

CCTGTTGGAT TTTAGTCTGT AACAAAGTTG TCGCTATTGT GCTTCATCCG TTCAGCTGAA

596

	CAGGAGGATG GATACAGCCG CGAGGCTAAA AAACGGATTT CCTCTTCCTA GCTTAAAATC	600
	TGATTTACAC TGTTTGTTT TTTAAGAAAC AAAAGTGCAT AGTTTAGATT TTTTTTTTTG	660
5	TTGAATATGT TTGTTCTTGG ACTTTATGAG AGAGTCTTAT AAGAATCACG ATTTTCTACA	720
	CCTGTCATTG AGCCAAGAAA GTCCAGTTTA TGACACGTAT GTACTAGTGA ACACCGTCCT	780
10	CGATCTGTAC GAAATGTGAA ATGTTTAGGG ACATCTCCAT GCTGTCACTT GTGATTTGCC	840
	CTCTTATGTA TTTTGGTCAT ATTGCCAACT GGAAAGTCAA AATTTTCTAA CAACTTTAAG	900
	TAAGTTCTTT GAAGACTTAG TGCTGTTTTT AATCCAGTTT AGAAAGTAAC TTAATTTTAA	960
15	TACCACTACT AAAAATTCGA AAATTTCTTC TTTAATCACA TTCAATATGG TTAAAAGAAC	1020
	AACACTAATT GACATTGCGT GGGCTTTTTC TCCCTTTGTT TAAAATGTCA TTTGTTGAGC	1080
20	AAGAGITGTA TAGTATTATC TACTTACTTG AGGCTGTTAA TTTTTCATTA CAGTGTTTTG	1140
	TAAATGTATC CACGAGACCA TGATGCATTG TTTTGTGCTC AACTTGTGTT TTGTATTTAA	1200
	AGCATTTGA ATGAAGTGTA TTTTATAAGC ATTTAATATT TATGCTCTTT AGAATGGAAC	1260
25	ACAGAAAACA AACCTTATAA GTCCTGATTA ATCTGAACCA ATAACCTGTG TGGCCTACAA	1320
	AGTATAATTC TATTAAATGT TCCTTAAAAC AAAAAAAAAA	1374
30		
	(2) INFORMATION FOR SEQ ID NO: 47:	
35	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 596 base pairs	
	(B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
ю	(C) STRANDEDNESS: double	
10	(C) STRANDEDNESS: double (D) TOPOLOGY: linear	60
-	(C) STRANDEDNESS: double (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47:	60 120
10 15	(C) STRANDEDNESS: double (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47: GAATTCGNCA CGAGATTACT TGGACATGAA AGAACTCAGG TTCAAGTTTA TTCATTTACT	
-	(C) STRANDEDNESS: double (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47: GAATTCGNCA CGAGATTACT TGGACATGAA AGAACTCAGG TTCAAGTTTA TTCATTTACT AAGTTAGTTA AATCATGGC CTTCCATGAG CCTTCATTTG GTAACTTCGA AAATGGAAAT	120
-	(C) STRANDEDNESS: double (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47: GAATTCGNCA CGAGATTACT TGGACATGAA AGAACTCAGG TTCAAGTTTA TTCATTTACT AAGTTAGTTA AATCATGTGC CTTCCATGAG CCTTCATTTG GTAACTTGGA AAATGGAAAT AATAACACTA GTCATATATA TTCTACACTG CTACCATATG GACCAAAGGG ATTATAGATT	120 180
15	(C) STRANDEDNESS: double (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47: GAATTCGNCA CGAGATTACT TGGACATGAA AGAACTCAGG TTCAAGTITA TTCATTTACT AAGTTAGTTA AATCATGTGC CTTCCATGAG CCTTCATTTG GTAACTTGGA AAATGGAAAT AATAACACTA GTCATATATA TTCTACACTG CTACCATATG GACCAAAGGG ATTATAGATT ACAATCACCA TCATTCCTGC TGACAGGTAT ATAGAAAACA ATTTCATTGA AGAAAAGTCC	120 180 240
i5	(C) STRANDEDNESS: double (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47: GAATTCGNCA CGAGATTACT TGGACATGAA AGAACTCAGG TTCAAGTITA TTCATTTACT AAGTTAGTTA AATCATGTGC CTTCCATGAG CCTTCATTTG GTAACTTGGA AAATGGAAAT AATAACACTA GTCATATATA TTCTACACTG CTACCATATG GACCAAAGGG ATTATAGATT ACAATCACCA TCATTCCTGC TGACAGGTAT ATAGAAAACA ATTTCATTGA AGAAAAGTCC TTACATTTAT CCTTTTCCTA ATATCTGCAT GGGTAAACTA ATAAATATAG TCATTAGAAA	120 180 240 300
15	(C) STRANDEDNESS: double (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47: GAATTCENCA CGAGATTACT TEGACATGAA AGAACTCAGG TTCAAGTITA TTCATTTACT AACTTAGTTA AATCATGTGC CTTCCATGAG CCTTCATTTG GTAACTTGGA AAATGGAAAT AATAACACTA GTCATATATA TTCTACACTG CTACCATATG GACCAAAGGG ATTATAGATT ACAATCACCA TCATTCCTGC TGACAGGTAT ATAGAAAACA ATTTCATTGA AGAAAAGTCC TTACATTTAT CCTTTTCCTA ATATCTGCAT GGGTAAACTA ATAAATATAG TCATTAGAAA ACCCTTATTA TTATTATTAG TTCAATGTGA GAACTGCTGC AGAAAAAATA TGCTTTATAA	120 180 240 300 360

GACCCTGTCT TAATTAAAAA AAAAAAAAAA AAACTCGAGG GGGGCCCGGT ACCCTA

5	(2) INFORMATION FOR SEQ ID NO: 48:	
10	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 851 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 48:	
15	CACATGAAGA CACACAGTGG TGAGAAGCCC TTCCGCTGCG CCCGCTGTCC TTATGCCTCT	60
	CCTCATCTGG ATAACCTGAA ACGGCACCAG CGCGTCCATA CAGGAGAGAA GCCCTACAAG	120
20	TGCCCCCTCT GCCCTTATGC CTGTGGCAAT CTGGCCAACC TCAAGCGTCA TGGTCGCATC	180
20	CACTCTGGTG ACAAACCTTT TCGGTGTAGC CTTTGCAACT ACAGCTGCAA CCAGAGCATG	240
	AACCTCAAAC GTCACATGCT GCGGCACACA GGCGAGAAGC CTTCCGCTGT GCCACCTGCG	300
25	CCTATACCAC GGGCCACTGG GACAACTACA AGCGCCACCA GAAGGTGCAT GGCCACGGTG	360
	GGGCAGGAGG GCCTGGTCTC TCTGCCTCTG AGGGCTGGGC CCCACCTCAT AGCCCACCCT	420
30	CTGTTTTGAG CTCTCGGGGC CCACCAGCCC TGGGGACTGC TGGCAGCCGG GCTGTCCACA	480
30	CAGACTCATC CTGAACTAGG TCCTTCTTCC CCATGTTTTA TACAGACGGA CCAGAAGCCA	540
	CCTTTTTCTC CCCCGCTGGC CAGGGGCTCC ACACAGACTA ACCTAGGCAC TATAAGGACC	600
35	AGCCCAACCC CATGGGCGGG GGGGCCCATA TGGACCAGGG GACCTTGCCT TGACTGAGGC	660
	ACTICACGAG CICAGIGAGA AGGGCCCTGT ATTCACCTCC ACTGCCCCCA GGGGCTGTGG	720
40	ACAAACCGGC TGGGGGACTG CCCAGCCTCC CACCTGTTTA TTTAACTTAT TTCAGTGCTT	780
40	TATAATAAAG GAAACACTAA CAAAGCCATG TCTATGCTGA ATTGGCAATG GCAGGCAATT	840
	TGGCCTTACC C	851
45		
	(2) ANDORROW DOD, GDO, TO, NO. 40.	
50	(2) INFORMATION FOR SEQ ID NO: 49:	
50	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 2020 base pairs	
	(B) TYPE: nucleic acid (C) STRANDEDNESS: double	
55	(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 49:	
60	GTGAAATGAA AACAGTCTTT TTATAGCCTT TAGCTTGTGA GTTTGGAAGT TTGGGGGGTC	60
α	\mathbf{m}	120

	ACACAGACCC	AGGTGAACAC	GCTGACTGTG	AACCTGCCCT	GTATCCGGAG	CTCTCCTCGG	180
5	CACTGAGGGG	ATGCAACAAA	ATTAGGAGAG	GWTCCTTGCT	CCCAACGICT	ACTICTCCTA	. 240
	CCTCAACAGG	GGTCCAGGGT	GCAGTGAACT	CAGTTCTTGG	CCCTTGGGTG	AGGAȚTCATG	300
	GATGAATGAA	AGCTAGACCT	GATGGGGAGG	CATTATGACT	AAATAGGCCC	AGCCTCCTTC	360
10	CCTTCCAGCT	CTGTCCTAGG	AGCATAGGCG	GGAAATCTGA	GTAGAGTCTG	ACTGCAGTTT	420
	TTGCTTATGA	TTTGTAAAAG	CCGTCATGGG	GTCAATAAGA	AAATAGGGGT	GATGGAGGG	480
15	GAGAAGCCCA	GGACTGGGAG	AATCGCACGT	GCCCCAGGGG	TTTTCACCAA	GGATTTTCAA	540
	GACAAACTGG	AGTAAGAATT	AAAGCCCCAG	AGGATTTAAT	TATCCTGGTT	TGCAAAAGAG	600
	CCTCCCATGC	CAGTACCGCC	CAGCCTTGGA	GGCCGGAATG	CTCATGGCCC	CIGIGGICIG	660
20	CTTGTCCTTC	AGCCCATGCC	CAGCAGATAC	CTCTCTGACT	GGAGACGGGC	TCAAAGCTGG	720
	ATTAGAAAGG	GGAGMGGCAC	TTGTGACTTT	GTTTGACTCT	GTGACTCACT	TCCTCGCTCA	780
25	CACCTTGTTT	GAACTACTGG	ACTITCAACT	GGCTTTCCTT	AGGTCAGGCA	AGCAGACAGC	840
	TCCCCACTGA	AGAGGTCTGT	ACAGTGACAA	ccccccccc	CAGCAAGGAC	ACAGATGCAG	900
	CCACAGTAAG	GCTCCATCAG	GACTGGGTCA	GTGATGGCAA	CAGGATGGCC	AAGGATGGCT	960
30	CTAGAACAYT	CTGTCCATGC	GTCACTCCCC	CCAGTTTTRT	TTTTAGCTTT	GGCTTCAGGG	1020
	AGTGACAGCC	ATCACAAATA	GCCACATTCT	GCTCTACTCT	CCAACATACC	AGATTSTACA	1080
35	CTGTTGTTAT	TTCATGAGAC	GTGAATGTTG	CAGAGAGTGG	GGGGATTCTG	GTTGTTAAGĢ	1140
	AACTTACACT	GGGGAGCTTT	ACTCTTCCGT	GTCAACAATG	TGACTACATG	TTCTCCAGAT	1200
	TAGCCACACA	TGCAAACATC	AGTGTCCTTC	TAGCTTTANC	CGAGAAAGAA	ACCAGTCCCA	1260
40	GGGAATGAAT	GGTGGTCTCC	CCACTCCCGG	CAGCACTTTA	GGCAGCCCAT	AAGCTATGCG	1320
	AGAATGTGAA	CGCTCACCTT	GCTCCGTCAC	GCTTCTGACC	TACCACATAA	ACAGGAAGAA	1380
45	GCCAGTGACC	GGAACAGCTC	TAGGAATAAC	AAGTCAGAAT	AGAAGTGTCC	TTTATATTAC	1440
	CAGAAAATAT	GGGCTTGGCC	TAAGTCGCTG	TCTCCTAACC	TGCCGGGGTC	ATTCCCCACC	1500
	AAACACCCCA	TACTAAGGAG	CCATGAGCCA	CCTGGACATT	CACCTTTTCT	TTGACCATCT	1560
50	GGAGTCTGGG	GCAACTTAAG	GAAGGCNCCA	CACAGTGGTG	CAGGCACATT	TCCAAGCGTA	1620
	GGTGTCCCTG	GCTTTTGTGG	CCAAAGCTAG	TGTTATGGTC	AACAACAGGC	CAGGGTCTGT	1680
55	GGGCACTGA	CCTTGAAAGT	GGCAAAATGG	AGGTTTCACA	GCTGTGCGG	GAGCAGGACG	1740
	GCTTGCTTCA	TCTAACAATC	TCAGTTTCCT	TTAAAAAAAG	AAAGAAAGGA	AAAGATTTCA	1800
	TAAGCAGGTG	TCAGTGGACA	GTTTAAGYAC	TTAACCATTT	CICITICITC	TTATGGATGT	1860
50	GAACTGTGCT	GTGGATAAAT	CATTTGTATT	TCTTGAATGT	TCTCTATGAC	TAACAGTTAT	1920

	TAAGTCGGTT GTGTATATGT GTAACTAATG TAACTGCCTT TTAAAATTTC ATTACAATAA	· 19 8
5	AAATGACTTT GCTCTGAAMA AAAAAAAAA AAAAACTCGA	2020
10	(2) INFORMATION FOR SEQ ID NO: 50: (i) SEQUENCE CHARACTERISTICS:	
15	(A) LENGTH: 2432 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 50:	
20	ATGAAGGTC GTTGGTGGGA AAGATGGCGG CGACTCTGGG ACCCCTTGGG TCGTGGCAGC	60
	AGTGGCGGCG ATGTTTGTCG GCTCGGGATG GGTCCAGGAT GTTACTCCTT CTTCTTTTGT	120
	TGGGGTCTGG GCAGGGGCCA CAGCAAGTCG GGGCGGGTCA AACGTTCGAG TACTTGAAAC	180
25	GGGAGCACTC GCTGTCGAAG CCCTACCAGG GTGTGGGCAC AGGCAGTTCC TCACTGTGGA	240
	ATCTGATGGG CAATGCCATG GTGATGACCC AGTATATCCG CCTTACCCCA GATATGCAAA	300
30	GTAAACAGGG TGCCTTGTGG AACCGGGTGC CATGTTTCCT GAGAGACTGG GAGTTGCAGG	360
50	TGCACTTCAA AATCCATGGA CAAGGAAAGA AGAATCTGCA TGGGGATGGC TTGGCAATCT	420
	GGTACACAAG GAATCGGATG CAGCCAGGGC CTGTGTTTGG AAACATGGAC AAATTTGTGG	480
35	GGCTGGGAGT ATTTGTAGAC ACCTACCCCA ATGAGGAGAA GCAGCAAGAG CGGGTATTCC	540
	CCTACATCTC AGCCATGGTG AACAACGGCT CCCTCAGCTA TGATCATGAG CGGGATGGGC	600
40	GGCCTACAGA GCTGGGAGGC TGCACAGCCA TTGTCCGCAA TCTTCATTAC GACACCTTCC	660
70	TGGTGATTCG CTACGTCAAG AGGCATTTGA CGATAATGAT GGCAAGCATG	720
	AGTGGAGGGA CTGCATTGAA GTGCCCGGAG TCCGCCTGCC CCGCGGCTAC TACTTCGGCA	780
45	CCTCCTCCAT CACTGGGGAT CTCTCAGATA ATCATGATGT CATTTCCTTG AAGTTGTTTG	840
	AACTGACAGT GGAGAGAACC CCAGAAGAGG AAAAGCTCCA TCGAGATGTG TTCTTGCCCT	900
50	CAGTGGACAA TATGAAGCTG CCTGAGATGA CAGCTCCACT GCCGCCCCTG AGTGGCCTGG	960
30	CCCTCTTCCT CATCGTCTTT TTCTCCCTGG TGTTTTCTGT ATTTGCCATA GTCATTGGTA	1020
	TCATACTCTA CAACAAATGG CAGGAACAGA GCCGAAAGCG CTTCTACTGA GCCCTCCTGC	1080
55	TGCCACCACT TTTGTGACTG TCACCCATGA GGTATGGAAG GAGCAGGCAC TGGCCTGAGC	1140
	ATGCAGCCTG GAGAGTGTTC TTGTCTCTAG CAGCTGGTTG GGGACTATAT TCTGTCACTG	1200
60	GAGTTTTGAA TGCAGGGACC CCGCATTCCC ATGGTTGTGC ATGGGGACAT CTAACTCTGG	1260

		4					
	TCTGGGAAGC	CACCCACCCC	AGGCAATGC	TGCTGTGATG	TGCCTTTCCC	TGCAGTCCTT	1320
	CCATGTGGGA	GCAGAGGTGT	GAAGAGAATT	TACGTGGTTG	TGATGCCAAA	ATCACAGAAC	1380
5	AGAATTTCAT	AGCCCAGGCT	GCCGTGTTGT	TTGACTCAGA	AGGCCCTTCT	ACTTCAGTTT	1440
	TGAATCCACA	AAGAATTAAA	AACTGGTAAC	ACCACAGGCT	TTCTGACCAT	CCATTCGTTG	1500
10	GGTTTTGCAT	TTGACCCAAC	CCTCTGCCTA	CCTGAGGAGC	TTTCTTTGGA	AACCAGGATG	1560
10	GAAACTTCTT	CCCTCCCTTA	CCTTCCTTTC	ACTCCATTCA	TIGICCICIC	TGTGTGCAAC	1620
	CTGAGCTGGG	AAAGGCATTT	GGATGCCTCT	CIGITGGGGC	CIGOGGCIGC	AGAACACACC	1680
15	TGCGTTTCAC	TGGCCTTCAT	TAGGTGGCCC	TAGGGAGATG	GCTTTCTGCT	TTGGATCACT	1740
	GTTCCCTAGC	ATGGGTCTTG	GGTCTATTGG	CATGTCCATG	GCCTTCCCAA	TCAAGTCTCT	1800
20	TCAGGCCCTC	AGTGAAGTTT	GGCTAAAGGT	TGGTGTAAAA	ATCAAGAGAA	GCCTGGAAGA	1860
20	CATCATGGAT	GCCATGGATT	AGCTGTGCAA	CTGACCAGCT	CCAGGTTTGA	TCAAACCAAA	1920
	AGCAACATTT	GTCATGTGGT	CTGACCATGT	GGAGATGTTT	CTGGACTTGC	TAGAÇCCTGC	1980
25	TTAGCTGCAT	GTTTTGTAGT	TACGATTTTT	GGAATCCCAC	TTTGAGTGCT	GAAAGTGTAA	2040
	GGAAGCTTTC	TTCTTACACC	TTGGGCTTGG	ATATTGCCCA	GAGAAGAAAT	TTGGCTTTTT	2100
30	TTTTCTTAAT	GGACAAGAGA	CAGTTGCTGT	TCTCATGTTC	CAAGTCTGAG	AGCAACAGAC	2160
<i>J</i> 0	CCTCATCATC	TGTGCCTGGA	AGAGTTCACT	GTCATTGAGC	AGCACAGCCT	GAGTGCTGGC	2220
	CTCTGTCAAC	CCTTATTCCA	CTGCCTTATT	TGACAAGGGG	TTACATGCTG	CTCACCTTAC	2280
35	TGCCCTGGGA	TTAAATCAGT	TACAGGCCAG	AGTCTCCTTG	GAGGCCTGG	AACTCTGAGT	2340
	CCTCCTATGA	ACCTCTGTAG	CCTAAATGAA	ATTCTTAAAA	TCACCGATGG	AACCAAAAAA	2400
40	AAAAAAAA	ААААААААА	AAAAAAAA	AA			2432
45	(2) INFORM	ATION FOR S	EQ ID NO: 5	1:			
	(i)	SEQUENCE C	HARACTERISI	ICS:			
			NGTH: 2340 h	-		-	
		(B) TY	PE: nucleic	acid			

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 51:

GACGCTGGGG GCGGGTGGGG GCGCGGGTA CCGGGCTGGA CGGCCGGCCG GCGCCCCCTC 60

ATTAGTATGC GGACGAAGCG GCGGGCTGCG CGGAGNGACG TCCCCTGCAG CCGCGGACCG 120

AGGCAGCGGC GGCACCTGCC GGCCGAGCAA TGCCAAGTGA GTACACCTAT GTRAAACTGA 180

60 GAAGTGATTG CTCGAGGCCT TCCCTGCAAT GGTACACCCG AGCTCAAAGC AAGATGAGAA 240

	GCCCAGCTT	GTTATTAAAA	GACATCCTCA	AATGTACATT	CCTTCTCTTT	GGAGTGTGGA	300
5	TCCTTTATAT	CCTCAAGTTA	AATTATACTA	CTGAAGAATG	TGACATGAAA	AAAATGCATT	360
J	ATGTGGACCC	TGACCATGTA	AAGAGAGCTC	AGAAATATGC	TCAGCAAGTC	TTGCAGAAGG	420
	AATGTCGTCC	CAAGTTTGCC	AAGACATCAA	TGGCGCTGTT	ATTTGAGCAC	AGGTATAGCG	480
10	TGGACTTACT	CCCTTTTGTG	CAGAAGGSCC	CCAAAGACAG	TGAAGCTGAG	TCCAAGTACG	540
	ATCCTCCTTT	TOGGTTCCGG	AAGTTCTCCA	GTAAAGTCCA	GACCCTCTTG	GAACTCTTGC	600
15	CAGAGCACGA	CCTCCCTGAA	CACTTGAAAG	CCAAGACCTG	TCGGCGCTGT	GIGGITATIG	660
1.5	GAAGCGGAGG	AATACTGCAC	GGATTAGAAC	TGGGCCACAC	CCTGAACCAG	TTCGATGTTG	720
	TGATAAGGTT	AAACAGTGCA	CCAGTTGAGG	GATATTCAGA	ACATGTTGGA	AATAAAACTA	780
20	CTATAAGGAT	GACTTATCCA	GAGGGCGCAC	CACTGTCTGA	CCTTGAATAT	TATTCCAATG	840
	ACTTATTTGT	TGCTGTTTTA	TTTAAGAGTG	TIGATTICAA	CIGGCTTCAA	GCAATGGTAA	900
25	AAAAGGAAAC	CCIGCCATTC	TGGGTACGAC	TCTTCTTTTG	GAAGCAGGTG	GCAGAAAAA	960
45	TCCCACTGCA	GCCAAAACAT	TTCAGGATTT	TGAATCCAGT	TATCATCAAA	GAGACTGCCT	1020
	TTGRACATCC	TTCAGTACTC	AGAGCCTCAG	TCAAGGTTCT	GGGGGCCGAG	ATAAGAACGT	1080
30	CCCCACAATC	GGTGTCATTG	CCGITGTCTT	AGCCACACAT	CTGTGCGATG	AAGTCAGTTT	1140
	GGCGGGTTTT	GGATATGACC	TCAATCAACC	CAGAACACCT	TTGCACTACT	TCGACAGTCA	1200
35	ATGCATGGCT	GCTATGAACT	TTCAGACCAT	GCATAATGTG	ACAACGGAAA	CCAAGTTCCT	1260
	CTTAAAGCTG	GTCAAAGAGG	GAGTGGTGAA	AGATCTCAGT	GGAGGCATTG	ATCGTGAATT	1320
	TTGAACACAG	AAAACCTCAG	TTGAAAATGC	AACTCTAACT	CTGAGAGCTG	TTTTTGACAG	1380
40	CCTTCTTGAT	GTATTTCTCC	ATCCTGCAGA	TACTTTGAAG	TGCAGCTCAT	GTTTTTAACT	1440
	TTTAATTTAA	AAACACAAAA	AAAATTTTAG	CTCTTCCCAC	TTTTTTTTC	CTATTTATTT	1500
45	GAGGTCAGTG	TTTGTTTTTG	CACACCATTT	TGTAAATGAA	ACTTAAGAAT	TGAATTGGAA	1560
,,	AGACTTCTCA	AAGAGAATTG	TATGTAACGA	TGTTGTWTTG	ATTTTTAAGA	AAGTAATTTA	1620
	ATTTGTAAAA	CTTCTGCTCG	TTTACACTGC	ACATTGAATA	CAGGTAACTA	ATTOGAAGGA	1680
50	GAGGGGAGGT	CACTCTTTTG	ATGGTGGCCC	TGAACCTCAT	TCTGGTTCCC	TGCTGCGCTG	1740
	CTTGGTGTGA	CCCACGGAGG	ATCCACTCCC	AGGATGACGT	GCTCCGTAGC	TCTGCTGCTG	1800
55	ATACTGGGTC	TGCGATGCAG	CGGCGTGAGG	CCTGGGCTGG	TTGGAGAAGG	TCACAACCCT	1860
33	TCTCTGTTGG	TCTGCCTTCT	GCTGAAAGAC	TCGAGAACCA	ACCAGGGAAG	CTGTCCTGGA	1920
	GGTCCCTGGT	CGGAGAGGGA	CATAGAATCT	GTGACCTCTG	ACAACTGTGA	AGCCACCCTG	1980
60	GGCTACAGAA	ACCACAGTCT	TCCCAGCAAT	TATTACAATT	CTTGAATTCC	TTGGGGATTT	2040

60

	•	
	TTTACTGCCC TTTCAAAGCA CTTAAGTGTT AGATCTAACG TGTTCCAGTG TCTGTCTGAG	2100
5	GTGACTTAAA AAATCAGAAC AAAACTTCTA TTATCCAGAG TCATGGGAGA GTACACCCTT	2160
	TCCAGGAATA ATGITTTGGG AAACACTGAA ATGAAATCTT CCCAGTATTA TAAATTGTGT	2220
	ATTTAAAAAA AAGAAACTTT TCTGAATGCC TACTGGCGGT GTATACCAGG CAGTGTGCCA	2280
10	СТТГАААААС АТСАААААСА АТАААААСТТ ТТСАССААМА АААААААААА	2340
15		
13	(2) INFORMATION FOR SEQ ID NO: 52:	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 601 base pairs (B) TYPE: nucleic acid	
20	(C) STRANDEDNESS: double	
	(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 52:	
25	AGTAGGGGAG ACTGAGACTG ACCGGTAGCC AGGCAGGCGG ACGACGCACG CCCGGACAGA	60
	CTGAGCAGGC GCCGGAGAAC CACTCACAGG TTCCCCCCGC CTTTCCCTTT GAAANCTAGG	100
	·	120
30	CTTTTGCCTT TCCCGTGGCG CCCGAGAGAG AATGCTGGAC TCTGCCGACT TCAGCGCAAC	. 180
	TAANGATITC TCAAGCTAGG GGACAAACGA TCAGCCCAAT CCTGAGAAGG GCCGAACCAA	240
	GCACCCCGTC CCCATCCCCC TCCCCTCCCC CGACTAAACT CGGGCGCCAA ACCCAGCCCT	300
35	TCTCTAACCA CCCTACTTCC TCCTCTCCTT TCTAGCATGG TGGCTGTATG GACAGTCTGA	360
	CAGAACAGAG ACTGACATCT CCCAATCTGC CGGCCCCCCA CCTGGAACAC TACAGTGTTC	420
10	TGCATTGCAC CATGACCCTG GATGTGCAAA CTGTAGTCGT TTTTGCCGTG ATTGTAGTCC	480
•0	TCCTGCTTGT CAATGTCATA CTCATGTTTT TCCTGGGAAC GCGCTGAATG GAGTCCAGNC	540
	ACCTGAGCTG TCGCGAACTC TCGCTTTGAT TTCATCCCGA GAGCCACCGA GAAGAAAAAA	600
1 5	A	601
50	(2) INFORMATION FOR SEQ ID NO: 53:	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 359 base pairs	
55	(B) TYPE: nucleic acid (C) STRANDEDNESS: double	
_	(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 53:	

CTCGTGCCGA ATTCGGCACG AGAGATGGTA CTTTTAAGAG GTAATTAGGT TGCTAAGATG

	GATTAACATC TITCTCTTGA CACTGAGACT GGGTTCTCCT GGGAATGGTT AGTTCCCAAG	120
5	AGAGIGAGIT GITATAAAAC AATGCTGCCT CTTCTATTIT GCGCTTTTIG TTTGCACAAA	180
3	CTCGGTCCCC TTCTGTTTCT CTACGATGTT TTGATGCRGC ATGAGGCAGT CATGAGAACC	240
	CACCAGATAC AGCTGCCTGA TCCTGAATTT CCCAGCCAAC AGAACCAAGT GCTAAATAAA	300
10	ACTICITITTA ATRAGITARA ARRARARAR ARRARARA ARRARARARA	359
15	(2) INFORMATION FOR SEQ ID NO: 54:	•
20	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1141 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double	
	(b) Toronogi. Illiear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 54:	
25	GGCACGAGCT GCTGAGGCGT GAGAATGGCG TCCCGCGGCC GGCGTCCGGA GCATGGCGGA	60
	CCCCCAGAGC TGTTTTATGA CGAGACAGAA GCCCGGAAAT ACGTTCGCAA CTCACGGATG	120
30	ATTGATATCC AGACCAGGAT GGCTGGGGGA GCATTGGAGC TTCTTTATCT GCCAGAGAAT	180
	AAGCCCTGTT ACCTGCTGGA TATTGGCTGT GGCACTGGGC TGAGTGGAAG TTATCTGTCA	240
	GATGAAGGCC ACTATTGGGT GGGCCTGGAT ATCAGCCCTG CCATGCTGGA TGAGGCTGTG	300
35	GACCGAGAGA TAGAGGGAGA CCTGCTGCTG GGGGATATGG GCCAGGGCAT CCCATTCAAG	360
	CCAGGCACAT TIGATGGITG CATCAGCATT TCTGCTGTGC AGTGGCTCTG TAATGCTAAC	420
40	AAGAAGTCTG AAAACCCTGC CAAGCGCCTG TACTGCTTTT TTGCTTCTCT TTTTTCTGTT	480
	CTCGTCCGGG GATCCCGAGC TGTCCTGCAG CTGTACCCTG AGAACTCAGA GCAGTTGGAG	540
	CTGATCACAA CCCAGGCCAC AAAGGCAGGC TTCTCCGGTG GCATGGTGGT AGACTACCCT	600
45	AACAGIGCCA AAGCAAAGAA ATICTACCIC TGCITGITIT CIGGGCCTIC GACCITTATA	660
	CCAGAGGGGC TGAGTGAAAA TCAGGATGAA GTTGAACCCA GGGAGTCTGT GTTCACCAAT	720
50	GAGAGGTTCC CATTAAGGAT GTCGAGGCGG GGAATGGTGA GGAAGAGTCG GGCATGGGTG	780
50	CTGGAGAAGA AGGAGCGGCA CAGGCGCCAG GGCAGGGAAG TCAGACCTGA CACCCAGTAC	840
	ACCGGCCGCA AGCGCAAGCC CCGCTTCTAA GTCACCACGC GGTTCTGGAA AGGCACTTGC	900
55	CTCTGCACTT TTCTATATTG TTCAGCTGAC AAAGTAGTAT TTTAGAAAAG TTCTAAAGTT	960
	ATAAAAATGT TITCIGCAGT AAAAAAAAG TICTCIGGGC CGGGCGIGGI GGCTCACACC	1020
60	TGTAATCCCA GCACCTTGGG AGGCTGAGGT GGGAGGATCA TTTGAGGCCA GGAGTTTGAG	1080
60		

	ACCIGCUIGG GCAACATAAT GAAACTTCCT TTCCAGGGAG AAAAAAAAAA	1140
	A	1141
5	. '	
	(2) INFORMATION FOR SEQ ID NO: 55:	
10 15	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1560 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	1
13	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 55:	
	TCCTTCTCTG GGGCGGTCGC GTTGGCAGCG GATGCGGGAA GCCGGACTCT GGGCGTCATG	60
20	TACTACAAGT TTAGTGGCTT CACGCAGAAG TTGGCAGGAG CATGGGCTTC GGAGGCCTAT	120
	AGCCCGCAGA TINAAAGCCT GTGGTTTCCA CAGAAGCACC ACCTATCATA TITGCCACAC	180
25	CAACTAAACT GACCTCCGAT TCCACAGTGT ATGATTATGC TGGGAAAAAC AAAGTTCCAG	240
23	AGCTACAAAA GTTTTTCCAG AAAGCTGATG GTGTGCCCGT CTACCTGAAA CGAGGCCTGC	300
	CTGACCAAAT GCTTTACCGG ACCACCATGG CGCTGACTGT GGGAGGGACC ATCTACTGCC	360
30	TGATCGCCCT CTACATGGCT TCGCAGCCCA AAAACAAATG AGTTAGGCTG CAGAGGACTG	420
	GITTGITTIT TGGCATAAAC CCTTTGAAGT TCCTTTFICA, TTGITAAATT AAAATTTITT	480
35	TTTTTACTTG GATGCTTAA CATTITTGCA AGAAAAATAG GAAGATATGA AGATGATGTT	540
55	TIGGITIGIT TATGAAATGC ATATGGCTTG TCAGAGCTCA TTCGACAGTT AAAGCCATTG	600
	TTTAAAGAAA CGGTGCTTTG CTCTGTGTTT GTGCTCCTGA TTTCCCTGGA GGTTCTGGAT	660
40	GAAGGCTGAA CACAGGCTTG TTAATGTCAG TCTGTGCTGA GGACCTCAGG GACTTGAGGT	720
	TGCATTTTTG AGCATGGGGT GCAGGAGCCT TTCTGGATTT GGATGTGGCT ATGGAAAGAA	780
45	CACAGAAGCC AAGGTCATGT GCATGAAATG AGGAGTTTGA GTTAGTCACC TCGGGGATTT	840
+3	TTTCCATTTT GCAGTAAAAT GTTAAATTAA TGTAGCCTGC CTCTATTTGT TGGGCAGGTA	900
	ATTTCAAAGG GTTATTTGCC TCATCTCCTA TCTTTAGTGA AATCTTATGT GTAATTGTGT	960
50	GTATTTATTC CACCGTGGGA ACAGAGAATA CCTGTTTAGT GTTGCACTTT AGACTGGTGT	1020
	CIGITITGIT AATGCAGCIG IGCCACAAAT ICTCCTITAT CITITAAAAA IGITATAGCI	1080
	TTAAATTTTG ATTTATTTTG ACTGTGGAAT AAATACATGA ATGAAAAATT TTAAGTTTGA	1140
55	AGTICITIGA ATGACCITIC AGAGTAATIT CAGAACACCA GCAGCATCIT AAACCIGAGT	1200
	CTAATTTCTT TCTTGTTAAT TAGGCACCAG ATAATCTTTA TAAAATGGTC TTAAAAGCTA	1260
50	GTAATAGGAG CTTAATGGCA ATKGATGATT ACCACAKGGT TTTTTATAAA AACCTGCCTG	1320

	CCCCTWAGTG	AAAGGTACCT	GTAACYCACA	GTYCATTTAG	ACACTAATTT	CCTYTGCYGT	1380
5	CATGATTGGK	AGACTTCACT	TACCCTATAT	TAATTTTGAA	AAAAGGTGGA	ATTTTATTAT	1440
	ATATGAAGGA	ATAGTTTGTA	TCTTACCATA	GCACAGAACA	GTGACCTCTT	GCTCAGGATA	1500
	AGATGTGGTG	ATTTGAAAAT	ACTCATAGTA	GCCTTGCAGT	GATACCTCTC	TCNCTCTCTC	. 1560
10		i. 1		,		ı	
	(2) INFORM	ATION FOR SI	70 TD NO. 54		•		
15				,			
13	(1)		GTH: 1507 b	ase'pairs '			
			E: nucleic ANDEDNESS:			•	
20			OLOGY: line		·		
20	(xi) SEQUENCE	DESCRIPTION	: SEQ:ID NO	: 56:		
	GGAACGCAGA	GCGGAGCGTG	GAGAGCGGAG	CGAAGCTGGA	TAACAGGGGA	CCGATGATGT	60
25	GGCGACCATC	AGTTCTGCTG	CTTCTGTTGC	TACTGAGGCA	CGGGGCCCAG	GGGAAGCCAT	120
	CCCCAGACGC	AGGCCCTCAT	GGCCAGGGGA	GGGTGCACCA	GCCGCCCCC	CTGAGCGACG	180
30	CTCCCCATGA	TGACGCCCAC	GGGAACTTCC	AGTACGACCA	TGAGGCTTTC	CTGGGACGGG	240
	AAGTGGCCAA	GGAATTCGAC	CAACTCACCC	CAGAGGAAAG	CCAGGCCCGT	CTGGGGGGGA	300
	TCGTGGACCG	CATGGACCGC	GCGGGGGACG	GCGACGGCTG	GCIGICGCIG	GCCGAGCTTC	360
35	GCGCGTGGAT	CGCGCACACG	CAGCAGCGGC	ACATACGGGA	CTCGGTGAGC	GCGGCCTGGG	420
	ACACGTACGA	CACGGACCGC	GACGGCGTG	TGGGTTGGGA	GGAGCTGCGC	AACGCCACCT	480
40	ATGGCCACTA	CGCGCCCGGT	GAAGAATTTC	ATGACGTGGA	GGATGCAGAG	ACCTACAAAA	540
	AGATGCTGGC	TCGGGACGAG	CGGCGTTTCC	GGGTGGCCGA	CCAGGATGGG	GACTCGATGG	600
	CCACTCGAGA	GGAGCTGACA	GCCTTCCTGC	ACCCCGAGGA	GTTCCCTCAC	ATGCGGGACA	660
45	TCGTGATTGC	TGAAACCCTG	GAGGACCTGG	ACAGAAACAA	AGATGGCTAT	GTCCAGGTGG	720
	AGGAGTACAT	CGCGGATCTG	TACTCAGCCG	AGCCTGGGGA	GGAGGAGCCG	GCGTGGGTGC	780
50	AGACGGAGAG	GCAGCAGTTC	CGGGACTTCC	GGGATCTGAA	CAAGGATGGG	CACCTGGATG	840
	GGAGTGAGGT	GGGCCACTGG	GTGCTGCCCC	CTGCCCAGGA	CCAGCCCCTG	GTGGAAGCCA	900
	ACCACCTGCT	GCACGARAGC	GACACGGACA	AGGAYGGGCG	GCTGAGCAAA	GCGSAAATCC	960
55	TGGGTAATTG	GAACATGTTT	GTGGGCAGTC	AGGCCACCAA	CTATGGYGAG	GACCTGACCC	1020
	GGCACCACGA	TGAGCTGTGA	GCMCCGNGCA	CCTGCCACAG	CCTCAGAGGC	CCGCACAATG	1080
60	ACCGGAGGAG	GGGCCGCTGT	GGTCTGGCCC	CCTCCCTGTC	CAGGCCCCGC	AGGAGGCAGA	1140

	TGCAGTCCCA GGCATCCTCC TKCCCCTGGG CTCTCAGGGA CCCCCTGGGT CGGCTTCTGT	1200
	CCCTGTCACA CCCCCAACCC CAGGGAGGGG CTGTCATAGT CCCAGAGGAT AAGCAATACC	1260
5	TATTTCTGAC TGAGTCTCCC AGCCCAGACC CAGGGACCCT NGGCCCCAAG CTCAGCTCTA	1320
	AGAACCGCCC CAACCCCTCC AGCTCCAAAT CTGAGCCTCC ACCACATAGA CTGAAACTCC	1380
10	CCTGGCCCCA GCCCTCTCCT GCCTGGCCTG GCCTGGGACA CCTCCTCTCT GCCAGGAGGC	1440
10	AATAAAAGCC AGCGCCGGGA AAAAAAAAA AAAAAAAAA AAAAAAAA	1500
	AAAAAAN	1507
15	· · ·	
	(2) INFORMATION FOR SEQ ID NO: 57:	•
20	(i) SEQUENCE CHARACTERISTICS:	
20	(A) LENGTH: 450 base pairs (B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double (D) TOPOLOGY: linear	
25	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 57:	
	GAATTCGGCA CGAGCAGTGT CCAACACTGT AGCTGGTGCC TGCCAGGTTC CCAGTGGCTG	50
30	GGGTCACCAG GTCTGAAGAG AGATGTGCTG GCTGCGGGCA TGGGSCCAGA TCYTCCTGCC	60
	AGTITICYTC TCCYTCTITC TCATCCAATT GCTTATCAGC TTCTCAGAGA ATGGTTTTAT	120 180
	CCACAGCCCC AGGAACAATC AGAAACCAAG AGATGGGAAT RAAGAGGAAT GTGCTGTAAA	240
35	GAAGAGTTGT CAATTGTGCA CAGAAGATAA GAAATATATG ATGAATAGAT AATTGAAAAG	300
	AGATCCTCCA GAAAGAGCAG AAGGAAGTTT CTTCAATGGC TTCCTTCAGG ATTTTAATCA	360
40	TCCTTACAGC CTCTTTGAGA ATGATTGAAC TTCCAAATTC CCTGAAGTTA AAATTTTAAA	420
	TTCTATTAAA CATTTTTCG AGTAAAAAA	450
		430
45		
	(2) INFORMATION FOR SEQ ID NO: 58:	
50	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1147 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double (D) TOPOLOGY: linear	
55	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 58:	
	GGCACGAGAC CCATTGAGCA GAAGGAGGCC AGGTGGGAAA GCTCCTGGGA AGAGCAGCCA	60
60	GACTOGACAC TOGGCTGCTT GAGTCCTGAG TCACAATTCA GAATTCCTGG GCTCCCTGGG	120
		

	TGCATTCTAT	CATTCCAGTT	GAAAGTTTGC	TICCTTCCAG	TCATGTGGCT	CTTCATTCTA	180
	CTCTCCTTGG	CTCTCATTTC	AGATGCCATG	GTCATGGATG	AAAAGGTCAA	GAGAAGCTTT	240
5	GTGCTGGACA	CGGCTTCTGC	CATCTGCAAC	TACAATGCCC	ACTACAAGAA	TCACCCCAAA	300
	TACTGGTGCC	GAGGCTATTT	CCGTGACTAC	TGCAACATCA	TCCCCTTCTC	CCCTAACAGC	360
10	ACCAATCATG	TGGCCCTGAA	GGACACAGGG	AACCAĞCTCA	TTGTCACTAT	GTCCTGCCTG	420
10	AACAAAGAAG	ACACGGGCTG	GTACTGGTGT	GCATCCAGC	GGGACTTTGC	CAGGGATGAC	480
	ATGGATTTTA	CAGAGCTGAT	TGTAACTGAC	GACAAAGGAA	CCTGGCCAAT	GACTITIGGTC	540
15	TGGGAAAGAC	TATCAGGCAC	AAAACCAGAA	GCTGCAAGGC	TCCCAAAGTT	GTCCGCAAGG	600
	CTGACCGCTC	_CAGGACGTCC.	ATTCTCATCA	TTTGCATACT	GATCACGGGT	TTGGGAATCA	660
20	TCTCTGTAAT	CAGTCATTTG	ACCAAAAGGA	GGAGAAGTCA	AAGGAATAGA	AGGGTAGGCA	720
20	ACACTTTGAA	GCCCTTCTCG	CGTGTCCTGA	CTCCAAAGGA	AATGGCTCCT	ACTGAACAGA	780
	TGTGACTGAA	GATTTTTTA	ATTTAGTTCA	TAAAGTGATG	CTACAACAGA	ATAATCACCA	840
25	TGACAACTGG	CCCCACACCT	CAGAGAÇTGA	TTCTGATCTC	CCAGGAATTC	TGAAGGTCCC	900
	TCTATCCTTG	ACAACAATCA	TTTGCAGCCA	GGTAGCAACG	GCAGTAGTCA	GAGGAGCTAT	960
30	GATAGACCAC	ACCCAAGCAA	GGCTGCCCTC	AAATAACATC	TCAAGATCTT	AGTTCTTATG	1020
50	CATTCCATCA	GTCAGAAGTG	AAGAAGAGGT	GGAGAATCTG	GATTGGGGAC	CAGGAAATCA	1080
	CTTGTATTTT	GTTAGCCAAT	AAATTCCTAG	CCAGIGITGA	ATGAAAAAAA	AAAAAAAA	1140
35	АААААА					·	1147
40			•				

40 (2) INFORMATION FOR SEQ ID NO: 59:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 777 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 59:

50 GECAGAGECT CCTCAGAAGG GCGTGGGCTC TCCAGTCTTC CACAGTCCCC ACCATGCCCT 60

GTTGCCTTAC CGCTGACGTA GCTCACCCAT CTTTTACTTG CCTGGCTAAG ATGCATGGCA 120

TYWCATTTCC TCCTTGTTGC ACTGCAGTCA GTCCCTCACT GCCCCCATCT CCTGGAAGAG 180

GAGCATAAGC TTTGCAAGGT CAGCCACTTC TCTGGGGTCA CACTAGTTAC ATCAAGACAG 240

GACTCCAGCT CATATGTGCC AGTGCAGACA CTCTTCATCC ACCTGGGCC CTGGGCTTGG 300

60 GACCTGGYTC CTTGCACAGC AGARGACCCG GAGGCTGAGA GGAGCTTGCG GTTGTGTCAT 360

	AGTCACCTGG	CCAGARGGAA	CGTGAGCCCC	TCCCAAGCTG	CAGARGGARG	GARCARGCGT	420
5	GGCTGTCAGC	ACCGAGGTAG	CAGAGAATTA	ACATTCTTGT	CAGCAGAGAA	TGAAGCAGGA	480
J	ATATAATTAA	AACTTTGCCC	TTGGAATAGC	TGATTCATTT	GAATTTTATT	CCACACGITT	540
	GAAAGAGGAA	AGAAAATGTG	AAGACTTGCA	GCCTGGTTCT	CGCCTGGCCT	GGGCTGGCCC	600
10	AGCTGTCAGG	CCCGGTTCCT	TTCTGAGCAT	TCAGTCCACT	GATGTTGACT	GAGGGCCAGG	660
	AGAGACCCTC	AGCAGGGTAT	TACCATATCA	GCCTCCTATC	GCTGCTGGGA	GAAATTACCA	720
15	TGAATTCAGT	GGCTTAAAAC	AACACACGAG	CCTCTCTGAG	CCTACCCTGG	CTCAGGA	777
1.5			ı'				

(2) INFORMATION FOR SEQ ID NO: 60:

20

25

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1191 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 60:

30	AAGANIGATI	TTCCTTACTC	TCCAAAGCGT	CAGCATITIG	AAGTTCTT	TATGAAAGIG	60
50	GGGGCAAGAA	TCAGGGTGAA	AATGAGTGTA	AACAAAGCCC	ATCCTGTGGT	CAGCACCCAC	120
	TGGAGGTGGC	CAGCAGAGTG	GCCTCAGATG	TTCCTGCACC	TGGCCCAGGA	GCCCAGGACA	180
35	GAGGTCAAAT	CTAGGCCCCT	TGGTCTGGCT	GGATTCATCA	GGCAAGATTC	GAAAACAAGA	240
	AAACCTCTAG	AACAAGAAAC	AATCATGTCT	GCAGCAGATA	CGGCACTGTG	GCCCTATGGC	300
40	CATGGCAATC	GTGAGCACCA	AGAGAATGAG	TTACAGAAAT	ATCTCCAATA	CAAAGACATG	360
40	CATCTCCTGG	ACAGTGGACA	GTCGCTGGGA	CACACACACA	CACTTCAAGG	CTCACACAAC	420
	CTAACAGCCT	TAAATATCTG	AAGAAACAGA	ATCACGACAT	TAAGTCAGCA	GAGGGAGAGG	480
45	TAGGCTGAAG	CAGCAGGAGG	CCAATTTTAT	ATCCCACAGA	TTTTTTTAAA	AATGACTCCC	540
	CAGCAAGGGG	TGGGGAGAAA	GCCACTGATT	TAGGAGAGTT	CTTGGCTCAG	CCAACCACTG	600
50	CGGTTATCTA	CACGITTTAC	AAAGGCACRG	AAGTAGAGAG	GGGCTGCACT	CACGACCCTC	660
50	CCCAGGCCC	GCACAGCCAG	ACACGGTGGG	TTCTTCCTTT	TTCCCTTCTG	GCCTTGGTGG	720
	AATTCCTACC	ACGGTGGCCT	CTGCCTTTGG	GACAATGCCT	TCATGCTCAT	CCCCGGGTCA	780
55	AGGATGGAGT	CTGTTACCAT	TTTCCAGGGG	AAATTCCAAG	GACCAGCCCC	GCCTCATTAC	840
	GTTCACCCCA	CAGGAAGGTG	ATCTGGAAAG	CCTGTAAACA	CGTACTCTCG	GTGGCTGAGT	900
60	GGTGTCACCA	AGCTGCTTTT	GTGCAGGGCT	GAAGCACAGA	. CAAGAGGGCA	GGCAGCTGCC	960

	GGAGGCCTGA AGTGGGGAGA GATCCCCGCA GGCCTGCAGG AGCCAGGGAG AACCTCCAAC	1020
	TOGATCTAAA CTGTGGGACA GCCCAGGCGT GCCCCTCTTC ACATGGCTCC CAGGCTCCCT	1080
5	CAAAGCCCTT CCCAGGCCCT GCAGGAAGAG AGGGAGGGTG AGGAGAGGCA GGGAGGGCAG	1140
	AGGTCGCCTG AAAGCCTGGG CTCCGAACTC CCTCAGCAGA GCTTTAAAGT G	1191
10		
10		
	(2) INFORMATION FOR SEQ ID NO: 61:	•
15	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1580 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	٠
20	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 61:	
	CCCCGCCCCC CGCCCACGAA GGAAGTGGCT GCTGCTCCGG CGCGGACCCA GAGCCGGTTC	60
25	GGCGCGTCGA CTGCCCAGAG TCCGCGGCCG GGCGCGGGAG GAGCCAAGCC GCCATGGCCT	120
	ACCACAGCTT CCTGGTGGAG CCCATCAGCT GCCACGCCTG GAACAAGGAC CGCACCCAGA	180
	TIGCCATCTG CCCCAACAAC CATGAGGIGC ATATCTATGA AAAGAGCGGT GCCAAATGGA	240
30	CCAAGGTGCA CGAGCTCAAG GAGCACAACG GGCAGGTGAC AGGCATCGAC TGGGCCCCCG	300
	AGAGTAACCG TATTGTGACC TGCGGCACAG ACCGCAACGC CTACGTGTGG ACGCTGAAGG	360
35	GCCGCACATG GAAGCCCACG CTGGTCATCC TGCGGATCAA CCGGGCTGCC CGCTGCGTGC	420
	GCTGGGCCCC CAACGAGAAC AAGITTGCTG TGGGCAGCGG CTCTCGTGTG ATCTCCATCT	480
	GTTATTTCGA GCAGGAGAAT GACTGGTGGG TTTGCAAGCA CATCAAGAAG CCCATCCGCT	540
40	CCACCGTCCT CAGCCTGGAC TGGCACCCCA ACAATGTGCT GCTGGCTGCC GGCTCCTGTG	600
	ACTICAAGIG TCGGATCITT TCAGCCTACA TCAAGGAGGT GGAGGAACGG CCGGCACCCA	660
45	CCCCGTGGGG CTCCAAGATG CCCTTTGGGG AACTGATGTT CGAATCCAGC AGTAGCTGCG	720
73	GCTGGGTACA TGGCGTCTGT TTCTCAGCCA GCGGGAGCCG CGTGGCCTGG GTAAGCCACG	780
	ACAGCACCGT CTGCCTGGCT GATGCCGACA AGAAGATGGC CGTCGCGACT CTGGCCTCTG	840
50	AAACACTACC ACTGCTGGCG CTGACCTTCA TCACAGACAA CAGCCTGGTG GCAGCGGGCC	900
	ACGACTGCTT CCCGGTGCTG TTCACCTATG ACGCCGCCGC GGGGATGCTG AGCTTCGGCG	960
55	GGCGGCTGGA CGTTCCTAAG CAGAGCTCGC AGCGTGGCTT GACGGCCCGC GAGCGCTTCC	1020
<i>J J</i>	AGAACCTGGA CAAGAAGGCG AGCTCCGAGG GTGGCACGGC TGCGGGCGCG GGCCTAGACT	1080
	CGCTGCACAA GAACAGCGTC AGCCAGATCT CGGTGCTCAG CGGCGGCAAG GCCAAGTGCT	1140
60	CGCAGTTCTG CACCACTGGC ATGGATGGCG GCATGAGTAT CTGGGATGTG AAGAGCTTGG	1200

AGTCAGCCTT GAAGGACCTC AAGATCAAAT GACCTGTGAG GAAGCCGCGG GAAGCCGGGA GAGGGGTCAG GGAGGCTAAT COCTGGGGTACC AATACGAGTT CCCATAGGGG CTGCTCCCTC AAGGGAGCTTTT CTTACCTATT CAAGGAATAC GTGCCTTTTT CTAACAAAAAAAAAA	GAACTGCTT CAAAATGTG 1320 AAAAGGGAG GGGACAGATG 1380 TTAAAATGCT TTCATTTATT 1440 GAACTGCTT CAAAATGTGG 1500
CTGGGGTACC AATACGAGTT CCCATAGGGG CTGCTCCCTC AAGGGAGCTTTT CTTACCTATT CAAGGAATAC GTGCCTTTTT CTGGAAAAAAAAAA	AAAAGGGAG GGGACAGATG 1380 TTAAATGCT TTCATTTATT 1440 GAACTGCTT CAAAATGTGG 1500 AATGACCCT CGCGATCTAG 1560
GGGAGCTTTT CTTACCTATT CAAGGAATAC GTGCCTTTTT CT GAAAAAAAAA AAAAATGCCC CCAAAGCACT ATGCTGGTCA' TC AGGTAATAAA ATGCAACTGT GTAAAAAAAA AAAAAAAAA AA AACTAGNCGG ACGCNTGGGT	TTAAATGCT TTCATTTATT 1440 GAACTGCTT CAAAATGTGG 1500 AATGACCCT CGCGATCTAG 1560
10 GAAAAAAAA AAAAATGCCC CCAAAGCACT ATGCTGGTCA' TO AGGTAATAAA ATGCAACTGT GTAAAAAAAA AAAAAAAAA AAAAAAAAAA	GAACTECTT CAAAATETEG 1500 AATGACCCT CECGATCTAG 1560
AGGTAATAAA ATGCAACTGT GTAAAAAAAA AAAAAAAAA AA	AATGACCCT CGCGATCTAG 1560
AACTAGNOGG ACGCNTGGGT	,
	1580
(2) INFORMATION FOR SEQ ID NO: 62:	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1117 base pairs	
(B) TYPE: nucleic acid (C) STRANDEDNESS: double	X
25 (D) TOPOLOGY: linear	1
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: (62:
GGCACGAGGC GCGATGCAGC ACAGGCTAGA GGCTGCGCAA SG	eccececc cecccitece 60
ACCCTCCGGG CCGGGCGGTT TGGCCCCTTA GCGCCCGGGC G1	TCGGGGCGG TAAAAGGCCG 120
GCAGAAGGGA GGCACTTGAG AAATGTCTTT CCTCCAGGAC CC	CAAGTTTCT TCACCATGGG 180
35 GATGTGGTCC ATTGGTGCAG GAGCCCTGGG GGCTGCTGCC TT	IGGCATTGC TGCTTGCCAA 240
CACAGACGTG TITCTGTCCA AGCCCCAGAA AGCGGCCCTG GA	AGTACCTGG AGGATATAGA 300
CCTGAAAACA CTGGAGAAGG AACCAAGGAC TTTCAAAGCA AA	AGGAGCTAT GGGAAAAAAA 360
TGGAGCTGTG ATTATGGCCG TGCGGAGGCC AGGCTGTTTC CT	ICTGTCGAG AGGAAGCTGC 420
GGATCTGTCC TCCCTGAAAA GCATGTTGGA CCAGCTGGGC GI	ICCCCCTCT ATGCAGTGGT 480
45 AAAGGAGCAC ATCAGGACTG AAGTGAAGGA TTTCCAGCCT TA	ATTTCAAAG GAGAAATCTT 540
CCTGGATGAA AAGAAAAAGT TCTATGGTCC ACAAAGGCGG AA	AGATGATGT TTATGGGATT 600
TATCCGTCTG GGAGTGTGGT ACAACTTCTT CCGAGCCTGG AA	ACGGAGGCT TCTCTGGAAA 660
CCTGGAAGGA GAAGGCTTCA TCCTTGGGGG AGTTTTCGTG GI	rgggatcag gaaagcaggg 720
CATTCTTCTT GAGCACCGAG AAAAAGAATT TGGAGACAAA GI	PARACCTAC TTTCTGTTCT 780
55 GGAAGCTGCT AAGATGATCA AACCACAGAC TTTGGCCTCA GA	AGAAAAAAT GATTGTGTGA 840
AACTGCCCAG CTCAGGGATA ACCAGGGACA TTCACCTGTG TT	CATGGGAT GTATTGTTTC 900
CACTCGTGTC CCTAAGGAGT GAGAAACCCA TTTATACTCT AC	CTCTCAGTA TGGATTATTA 960

	ATGTATTTA ATATTCTGTT TAGGCCCACT AAGGCAAAAT AGCCCCAAAA CAAGACTGAC	1020
	AAAAATCTGA AAAACTAATG AGGATTATTA AGCTAAAACC TGGGAAATAG GAGGCTTWAA	1080
5	ATGACTGCCM GCTGGTGCRT GCTCACACTT GGCCCAC	1117
10		1
10	(2) INFORMATION FOR SEQ ID NO: 63:	
15	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 361 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 63:	
20	CCCACGCGTG CKGGCGCCTG GCAGCCACCG CCTGGGAGGT TACTGTAAGG CCCGCAGCTC	60
	CCGCCAGCTC CCGCGGACTS CTGCCGCCTC CTTACCATGA AGCCAGTAAG TCGTCGCACG	120
25	CTGGACTGGA TITATTCAGT GTTGCTGCTT GCCATCGTTT TAATCTCCTG GGGCTGCATC	180
	ATCTATGCTT CGATGGTGTC TGCAAGACGA CAGCTAAGGA AGAAATACCC AGACAAAATC	240
	TTTGGGACGA ATGAAAATTT GTAACTCTTC TGGATTTAAT TATCTGAAAA TACAGTTCTT	300
30	TCCCTCATGC TTATGTAGAT ATAAAAATAA AATTCATAAT GCAAAAAAAA AAAAAAAAA	360
	G	361
35		
	(2) INFORMATION FOR SEQ ID NO: 64:	
40	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1668 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
45	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 64:	
	GGCACGAGGT CTGCCAAGCT ATAGACCATG GCTGTGAACA CATTTGTGTG AACAGTGACG	60
50	ACTCATACAC GTGCGAGTGC TTGGAGGGAT TCCGGCTCGC TGAGGATGGG AAACGCTGCC	120
50	GAAGAAGGAT GTCTGCAAAT CAACCCACCA TGGCTGCGAA CACATTTGTG TTAATAATGG	180
	GAATTCCTAC ATCTGCAAAT GCTCAKAGGG ATTTGTTCTA GCTGAGGACG GAAGACGGTG	240
55	CAAGAAATGC ACTGAAGGCC CAATTGACCT GGTCTTTGTG ATCGATGGAT CCAAGAGTCT	300
	TGGAGAAGAG AATTITGAGG TCGTGAAGCA GTTTGTCACT GGAATTATAG ATTCCTTGAC	360
60	AATTTCCCCC AAAGCCGCTC GAGTGGGGCT GCTCCAGTAT TCCACACAGG TCCACACAGA	420

	GTTCACTCTG	AGAAACTTCA	ACTCAGCCAA	AGACATGAAA	AAAGCCGTGG	CCCACATGAA	480
	ATACATGGGA	AAGGGCTCTA	TGACTGGGCT	GCCCTGAAA	CACATGTTTG	AGAGAAGTTT	· 540
5	TACCCAAGGA	GAAGGGGCCA	GCCCTTTCC	ACAAGGGTGC	CCAGAGCAGC	CATTGTGTTC	600
	ACCGACGGAC	GGGCTCAGGA	TGACGTCTCC	GAGTGGGCCA	GTAAAGCCAA	GGCCAATGGT	660
10	ATCACTATGT	ATGCTGTTGG	GGTAGGAAAA	GCCATTGAGG	AGGAACTACA	AGAGATTGCC	720
10	TCTGAGCCCA	CAAACAAGCA	TCTCTTCTAT	GCCGAAGACT	TCAGCACAAT	GGATGAGATA	780
	AGTGAAAAAC	TCAAGAAAGG	CATCTGTGAA	GCTCTAGAAG	ACTCCGATGG	AAGACAGGAC	840
15	TCTCCAGCAG	GGGAACTGCC	AAAAACGGTC	CAACAGCCAA	CAGTGCAACA	CAGATATCTG	900
	TTTGAAGAAG	ACAATCTTTT	ACGGTCTACA	CAAAAGCTTT	CCCATTCAAC	AAAACCTTCA	960
20	GGAAGCCCTT	TGGAAGAAAA	ACACGATCAA	TGCAAATGTG	AAAACCTTAT	AATGTTCCAG	1020
20	AACCTTGCAA	ACGAAGAAGT	AAGAAAATTA	ACACAGCGCT	TAGAAGAAAT	GACACAGAGA	1080
	ATGGAAGCCC	TGGAAAATCG	CCTGAGATAC	AGATGAAGAT	TAGAAATCGC	GACACATTTG	1140
25	TAGTCATTGT	ATCACGGATT	ACAATGAACG	CAGTGCAGAG	CCCCAAAGCT	CAGGCTATTG	1200
	TTAAATCAAT	AATGTTGTGA	AGTAAAACAA	TCAGTACTGA	GAAACCTGGT	TTGCCACAGA	1260
30	ACAAAGACAA	GAAGTATACA	CTAACTTGTA	TAAATTTATC	TAGGAAAAA	ATCCTTCAGA	1320
	ATTCTAAGAT	GAATTTACCA	GGTGAGAATG	AATAAGCTAT	GCAAGGTATT	TTGTAATATA	1380
	CTGTGGACAC	AACTTGCTTC	TGCCTCATCC	TGCCTTAGTG	TGCAATCTCA	TTTGACTATA	1440
35	CGATAAAGTT	TGCACAGTCT	TACTTCTGTA	GAACACTGGC	CATAGGAAAT	GCTGTTTTT [†]	1500
	TGTAYTGGAC	TTTACCTTGA	TATATGTATA	TGGATGTATG	CATAAAATCA	TAGGACATAT	1560
40	GTACTTGTGG	AACAAGTTGG	ATTTTTTATA	CAATATTAAA	ATTCACCACT	TCAGAGRAAA	1620
	АААААААА	АААААААА	AAAAAAAA	ааааааааа	AAANAAA		1668
15							

(2) INFORMATION FOR SEQ ID NO: 65:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1353 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 65:

55 GGGTCGACCC ACGCGTCCGC CCACGCGTCC GGATGGCTGC GCTGTTGCTG AGACACGTTG 60 GTCGTCATTG CCTCCGAGCC CACTTTAGCC CTCAGCTCTG TATCAGAAAT GCTGTTCCTT 120 60 TGGGAACCAC GGCCAAAGAA GAGATGGAGC GGTTCTGGAA TAAGAATATA GGTTCAAACC 180

	GTCCTCTGTC TCCCCACATT ACTATCTACA GTTGGTCTCT TCCCATGGCG ATGTCCATCT	- 240
5	GCCACCGTGG CACTGGTATT GCTTTGAGTG CAGGGGTCTC TCTTTTTTGGC ATGTCGGCCC	300
,	TGTTACTCCC TGGGAACTIT GAGTCTTATT TGGAACTTGT GAAGTCCCTG TGTCTGGGGC	360
	CAGCACTGAT CCACACAGCT AAGTTTGCAC TTGTCTTCCC TCTCATGTAT CATACCTGGA	420
10	ATGGGATCCG ACACTTGATG TGGGACCTAG GAAAAGGCCT GAAGATTCCC CAGCTATACC	480
	AGTCTGGAGT GGTTGTCCTG GTTCTTACTG TGTTGTCCTC TATGGGGCTG GCAGCCATGT	540
15	GAAGAAAGGA GGCTCCCAGC ATCATCTTCC TACACATTAT TACATTCACC CATCTTTCTG	600
13	TTTGTCATTC TTATCTCCAG CCTGGGAAAA GTTCTCCTTA TTTGTTTAGA TCCTTTTGTA	660
	TTTTCAGATC TCCTTGGAGC AGTAGAGTAC CTGGTAGACC ATAATAGTGG AAAAGGGTCT	720
20	AGTTTTCCCC TIGTTTCTAA AGATGAGGIG GCTGCAAAAA CTCCCCTTTT TTGCCCACAG	780
	CTTGCCTACT CTCGGCCTAG AAGCAGTTAT TCTCTCTCCA TATTGGGCTT TGATTTGTGC	840
25	TGAGGGTCAG CTTTTGGCTC CTTCTTCCTG AGACAGTGGA AACAATGCCA GCTCTGTGGC	900
	TTCTGCCCTG GGGATGGGCC GGGTTGGGG GTGGCTTTGG GTGCCACTGC	960
	CTGTGGGTTG CTGGCTTAAA GGACAATTCT CTTCATTGGT GAGAGCCCAG GCCATTAACA	1020
30	CCTACACAGT GTTATTGAAA GAAGAGAGGT GGGGGTGGAG GGGAATTAGT CTGTCCCAGC	1080
	TAGAGGGAGA TAAAGAGGGC TAGTTAGTTC TTGGAGCAGC TGCTTTTGAG GAGAAAATAT	1140
35	ATAGCTTTGG ACACGAGGAA GATCTAGAAA ATTATCATTG AACATATTAA TGGTTATTTC	1200
	TTTTTCTTGG ATTTCCAGAA AAGCCTCTTA ATTTTATGCT TTCTCATCGA AGTAATGTAC	1260
	CCTTTTTTC TGAAACTGAA TTAAATACTC ATTTTATCTT TGAAAAAAAA AAAAAAAACC	1320
40	TNGGGGGGG CCCCGGACCC NAATTGGCCC TAT	1353
		•
45	(2) INFORMATION FOR SEQ ID NO: 66:	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 1011 base pairs	
50	(B) TYPE: nucleic acid (C) STRANDEDNESS: double	
	(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 66:	
55	CGGAAGAAAG CAGCCATCCA GACATTTCAG AACACGTACC AGGTGTTAGC TGTGACCTTC	60
	AATGACACAA GTGATCAGAT TATTTCTGGT GGAATAGACA ATGATATCAA GGTCTGGGAC	120
60	TGCGCCAGAA CAAGCTAACC TACACCATGA GAGGCCATGC AGATTCAGTG ACTGGCCTGA	180

	GTTTAAGTTC	TGAAGGCTCT	TATCTTTTGT	CCAATGCAAT	GGACAATACA	GITCGIGICT	240
	GGGATGTCCG	GCCATTTGCC	CCCAAAGAGA	GATGTGTAAA	GATATTTCAA	GGAAATGTGC	300
5	ACAACTTTGA	AAAGAACCTT	CTGAGATGTT	CTTGGTCACC	TGATGGAAGC	AAAATAGCAG	360
	CTGGCTCAGC	CGACAGGTTT	GTTTATGTGT	GGGATACCAC	AAGCAGGAGA	ATATTGTATA	420
10	AGCTGCCCGG	CCATGCTGGC	TCCATCAATG	AAGTGGCTTT	CCACCCTGAT	GAGCCCATCA	480
	TTATCTCAGC	ATCGAGTGAC	AAGAGACTGT	ATATGGGAGA	GATTÇAGTGA	AGATATOGAC	540
	TGGAAGACTC	CAAGGCCGCT	TGTCTTTGAG	ACCTCAGACT	GCATAAGTGA	TGCCAAATGT	600
15	TGGATGTCCA	GGYTAGCACC	CTCCCTTCAG	ATGACCATTG	CTAGCAAGAA	ACAGGAGGCG	660
	GTGGCCATAT	TCCAAAAACC	ACTICIGICC	CATTTCACCA	GGATGACTAA	GGCAAGCTCC	720
20	CTGTGGCCTC	TAAAAACCAC	CTGCCAGATT	TCAGGGACTG	TTTTTTTTT	TCTTTTTCTT	780
-0	TTTTCCTGTT	TTCTAATGCA	GCCCAATGT	GACAAATTTG	TTGGTTGGGA	TTTTTTTTT	840
	TTTTTGTAAC	TGGCTTGTAT	GATATTTTCT	TTCTGTATTT	CTCTAȚATCA	TTTŢGTATTA	900
25	AAAGCCAAAT	AGATGCCTTT	TTACAAGARM	АААААААА	AAAAAAAA	AAAAAAA	960
	CTGGGAGGG	GGCCCGGTA	CCCAAATCGC	CGGATATGAT	CGTAAACAAT	С	1011

35

(2) INFORMATION FOR SEQ ID NO: 67:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1193 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 67:

	GGCCGGGCGG	TGCGCACTGC	GGGCGCATCC	CTGCCCCGGC	GCCGTCCGTG	CCCGCGGGAC	60
45	CTGACAGCCG	GGTCAGAGGG	CGAACTGTGC	TCAGGCCCGG	GCTGGACGCA	GAGCCAGAGC	120
45	TGTCCCCAGA	GGAGCAGAGG	GTCCTGGAAA	GGAAGCTGAA	AAAGGAACGG	AAGAAAGAGG	180
	AGAGGCAGCG	TCTGCGGGAG	GCAGGCCTTG	TGGCCCAGCA	CCCCCTCCC	AGGCGCTCGG	240
50	GGGCCGAACT	GCCTGGGAC	TACCTCTGCA	GATGGGCCCA	AAAGCACAAG	AACTGGAGGT	300
	TTCAGAAGAC	GAGGCAGACG	TGGCTCCTGC	TGCACATGTA	TGACAGTGAC	AAGGTTCCCG	360
55	ATGAGCACTT	CTCCACCCTG	CTGGCCTACC	TGGAGGGGCT	GCAGGGCCGG	GCCCGAGAGC	420
<i>55</i>	TGACGGTGCA	GAAGGCGGAA	GCCTGATGCG	GGAGCTGGAT	GAGGAGGGCT	CTGATCCCCC	480
	CCIGCCGGG	AGGGCCCAGC	GCATCCGACA	GNIGCIGCAG	CTGCTCTCCT	AGTGGGTTCA	540
60	GCGCGGGGCG	GGGCCGCTGC	CCAGTGCAGG	GCTGCCTCAG	ACCACACAGG	GIGCAGCTCC	600

	TCCGGCGGTG GGGGCCGGGT TCACCAGCAG GGCAGCGGCT GAGCAAGGGC TTTCAGCTCC	660
5	TCCGGTGGTG GGGCCCGGGA TCACCAGCAC CAGAGCCTCG CAAGGGCCCC TTCCCTCCTC	720
	CAGACCCTCC TTGGCCGGTG ACGCTGTGAC AGTGATGGCA GGTTCAGTGC CTTCAGCGCA	780
	GAGCGTGGAT GCTCTGGAAT CACCCGGACC CCTGGCCTTG GAGGGACCCT CCAGCCCCAG	, 840
10	GAATCTCCTT TCGACCGAAA TCTCTATTTT TCTACCGGGA ATATTTTAGA GATTCGCCCA	900
	TGCTGGCTCC TCCCGCCAGC TGCAAACCTG CACCTTCCGC CTGATTCCCG ATCCCCCTGC	960
15	GTGGCCCGCA TTCCTGGTCC CCTGCCTGCG TCCATCGAGG GGCCTGGCTG TGGCCTGTTT	1020
••	TCCTTTGACC CCACACAGCG TCATTGCGGG TCATGGGGGAG CCCCTGGTGG GAGCTTGTGG	1080
	AGTCGGATCA CGTACCTGTG CAGAAACCGC CTCTGTGGCT GCATTTGAAA TAAAACCCGA	1140
20	CCCAGCAGCA AAAAAAAAAA AAAAAANCNC NAGGGGGGGC CCGGNACCCA ATT	1193
25	(2) INFORMATION FOR SEO ID NO: 68:	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 560 base pairs	
30	(B) TYPE: nucleic acid (C) STRANDEDNESS: double	
	(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 68:	
35	GAATTCGGCA CGAGTTGGCA CATGATGCAA AATGCATTTC TCAGAGTAGA TTGCAGTCAA	60
	AAATGITGGA AACTACTAAG CATGIGCARA TAGCATGCAT GCTGCTGCTG ACCTGCCAGA	120
10	TATTTCTCCC TTCCTCCTT TCTCCCTCAT TTATTCATTC	180
	CATTAAAAA ATTATATGTA TGTTTTGTGC AAAGCACCCT ACTCAAGGCT GCGGGGTACA	240
	AAAGTATATC AGAAGCCTTG GGCTTTGACM WACTTCTCTG TAGTAGTGCT AGATTTGTGT	300
15	GGATCTGCCA CACTTACTCC AGGCCTCTTG TGACCTGTGC TTTGCATTAA TCTCTTAGGC	360
	TAAGCCACAT ACCTITICAT TATACAATCT TIGCTGATGC TAAGGACAGA TICCAAAGTG	420
50	CCCTCCTTAT AATTTTTGTA TITAATGCAA AGIGTAATCA AGAATAGGCC ATTGITAGGT	480
	CAATTGCTTT TCTGTATTTA TCTTTTCAAA CAATAAATAA TCAGTGGGAT GAAAAAGGGC	540
	CGGAAAAAA AAAAAAAAA	560

(2) INFORMATION FOR SEQ ID NO: 69:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1657 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: double(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 69:

	CGGACNGAGC	CCCCCCCCCC	CACTTCCTGT	GGAGGCCGCA	GCGGGTGCGG	GCGCCGACGG	60
10	GCGAGAGCCA	GCGAGCGAGC	GAGCGAGCCG	AGCCGAGCCT	CCCGCCGTCG	CCATGGGCCA	120
	GAACGACCTG	ATGGGCACGG	CCGAGGACTT	CGCCGACCAG	TTCCTCCGTG	TCACAAAGCA	180
15	GTACCTGCCC	CACGTGGCGC	GCCTCTGTCT	GATCAGCACC	TTCCTGGAGG	ACGGCATCCG	240
13	TATGTGGTTC	CAGTGGAGCG	AGCAGCGCGA	CTACATCGAC	ACCACCTGGA	ACTGCGGCTA	300
	CCTGCTGGCC	TCGTCCTTCG	TCTTCCTCAA	CTTGCTGGGA	CANTGACTGG	CTGCGTCCTG	360
20	GTGTTGAGCA	GGAACTTCGT	GCAGTACGCC	TGCTTCGGGC	TCTTTGGAAT	CATAGCTCTG	420
	CAGACGATTG	CCTACAGCAT	TTTATGGGAC	TTGAAGTTTT	TGATGAGGAA	ccreeccere	480
25	GGAGGAGGCC	TGTTGCTGCT	CCTAGCAGAA	TCCCGTTCTG	AAGGGAAGAG	CATGITTGCG	540
23	GGCGTCCCCA	CCATGCGTGA	GAGCTCCCCC	AAACAGTACA	TGCAGCTCGG	AGGCAGGGTC	600
	TIGCIGGITC	TGATGTTCAT	GACCCTCCTT	CACTTTGACG	CCAGCTTCTT	TTCTATTGTC	660
30	CAGAACATCG	TGGGGCACAG	CTCTGATGAT	TTTAGTGGCC	ATTGGTTTTA	AAACCAAGCT	720
	GGCTGCTTIG	ACTCTTGTTG	TGTGGCTCTT	TGCCATCAAC	GTATATTTCA	ACGCCTTCTG	780
35	GACCATTCCA	GTCTACAAGC	CCATGCATGA	CTTCCTGAAA	TACGACTTCT	TCCAGACCAT,	840
33	GTCGGTGATT	GGGGGCTTGC	TCCTGGTGGT	GCCCTGGGC	CCTGGGGGTG	TCTCCATGGA	900
	TGAGAAGAAG	AAGGAGTGGT	AACAGTCACA	GATCCCTACC	TGCCTGGCTA	AGACCCGTGG	960
40	CCGTCAAGGA	CTGGTTCGGG	GTGGATTCAA	CAAAACTGCC	AGCTTTTATG	TATCCTCTTC	1020
	CCTTCCCCTC	CCTTGGTAAA	GGCACAGATG	TTTTGAGAAC	TTTATTTGCA	GAGACACCTG	1080
45	AGAATCAATG	GCTTCAGGAC	ATGGGTTCTC	TTCTCCTGTG	ATCATTCAAG	TGCTCACTGC	1140
43	ATGAAGACTG	GCTTGTCTCA	GTGTTTCAAC	CTCACCAGGG	CIGICICITG	GTCCACACCT	1200
	CCCTCCCTCT	TAGTGCCGTA	TGACAGCCCC	CATCAAATGA	CCTTGGCCAA	GTCACGGTTT	1260
50	CTCTCTCGTC	AAGGTTGGTT	GGCTGATTGG	TGGAAAGTAG	GGTGGACCAA	AGGAGGCCAC	1320
	GTGAGCAGTC	AGCACCAGTT	CTGCACCAGC	AGCGCCTCCG	TCCTAGTGGG	TGTTCCTGTT	1380
55	TCTCCTGCCC	CTGGGTGGGC	TAGGGCCTGA	TTCGGGAAGA	TGCCTTTGCA	GGGAGGGGAG	1440
JJ	GATAAGTGGG	ATCTACCAAT	TGATTCTGGC	AAAACAATTT	CTAAGATTTT	TTTGCTTTAT	1500
	GTGGGAAACA	GATCTAAATC	TCATTTTATG	CTGTATTITA	TATCTTAGTT	GTGTTTGAAA	1560
60	ACGITTIGAT	TTTTGGAAAC	АСАТСААААТ	AAATAATGGC	GTTTGTTGTA	ААААААААА	1620

	ı	
	AAAAAAACTC GRGGGGGGC CCGGTACCCA AATCGCC	1657
· 5		
	(2) INFORMATION FOR SEQ ID NO: 70:	
10	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 711 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	,
15	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 70:	
	GGCACGAGCG AAGACCCTGT TCGGACCCTG CCCCGATTCC AGACTCAGGT AGATCGTCGG	60
20	CATACCCTCT ACCGTGGACA CCAGGCAGCC CTGGGGCTGA TGGAGAGAGA TCAGGTATCC	120
20	CCCAGGGAGT AGGGGCTACC TTGAGGGGAT GATAGACCTC CCCCACTCCC AGTGKKACTC	180
	TGGAAATATG AAGGAACTAG GGAGTGGAAG AGATTTCAGA GCTGGGGAGA GGAGTTCCTC	240
25	CCTTCAAAGC CAGCAACTGC CTTTGGGGAA TGTCGGGGGG TCTCTCCTTT CTCCTGCTTG	300
	TTTRAGGIGG TACACAGICC CCCCTTCAMC TGGSGGGAAG CTGINCCGGA CARACTCATC	360
20	TCAGCTTTCC CTTGGGGCAG GATCGGGGGC AGCAGCTCCA GCAGAAACAG CAGGATCTGG	420
30	AGCAGGAAGG CCTCGAGGCC ACACAGGGGC TGCTGGCCGG CGAGTGGGCC CCACCCCTCT	480
	GGRAGCTGGG CAGCCTCTTC CAGGCCTTCG TGAAGAGGGA GAGCCAGGCT TATGCGTAAG	540
35	CTTCATAGCT TCTGCTGGCC TGGGGTGGAC CCAGGACCCC TGGGGCCTGG GTGCCCTGAG	600
	TGGTGGTAAA GTGGAGCAAT CCCTTCACGC TCCTTGGCCA TGTTCTGAGC GGCCAGCTTG	660
40	GCCTTTGCCT TAATAAATGT GCTTTATTTT CAAAAAAAA AAAAAAAAAC T	711
45	(2) INFORMATION FOR SEQ ID NO: 71: (i) SEQUENCE CHARACTERISTICS:	
50	(A) LENGTH: 935 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: double(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 71:	
<i></i>	GGCACAGGGT GAAAGCCAGC TAAACCCCAA GTGGAGAAGT GAAAGACATG GTTGTTCCCA	60
55	TAAGTTTATT GCTCACATTA TGAAAGAAGC CATAGTCATG AGTGAACCAC TCCCTAGGTT	120
	GATAAGGAAA CCAACACGGA AGATCTCTTT CTGGAAGAAG CAGCCAGCCT CGTGAAGGAG	180
60	CGGCCCAGCC GCCGGGCCCG AGGGTCGCCT TTTGTTCGGA GTGGCACGAT TGTCCGTTCC	240

	CAGACATTCT CGCCTGGAGC ACGAAGCCAG TATGTTTGCA GACTTTATCG TAGTGACAGC	300
5	GACAGTTCAA CGCTGCCCCG GAAGTCCCCCC TTTGTCCGAA ATACTTTGGA AAGACGAACC	360
	CTTCGCTATA AGCAGTCATG CAGGTCTTCC CTGGCTGAGC TCATGGCCCG CACCTCCCTG	420
	GACTTGGAGC TGGATCTCCA GGCGTCGAGA ACACGGCAGA GGCAGCTGAA TGAGGAGCTC	480
10	TGCGCCCTCC GTGAGCTGCG GCAGCGGTTN GGAGGACGCC CAGCTCCGTG GCCAGACTGA	540
	CCTCCCACCC TGGGTGCTTC GGGACGAGCG GCTCCGTGGC CTGCTGCGGG AGCCGAGCGG	600
15	CAGACAAGAC AGACCAAACT TGACTACCGT CATGAGCAGG CGGCTGAGAA GATGCTGAAG	660
-0	AAGGCCTCCA AGGAGATCTA CCAGCTGCGT GGCAGAGCCA CAAAGAGCCC ATCCAAGTGC	720
	AGACCTTTAG GGAGAAGATA GCATTCTTCA CAAGGCCAAG GATCAACATA CCTCCTCTCC	780
20	CAGCCGACGA CGTCTGATGG AGTGCATTGT GCACATGAAG TATTTATCCA CCTGTTTTAT	840
	TTTCATGAAG TTCTTAGACT AGCTGAATTT GTCTTTAAAA TATTTGTGCA AAGCTATTAA	900
25	TATACACATT TTGTAAAAA AAAAAAAAA AAACT	935
	·	1
	(2) INFORMATION FOR SEQ ID NO: 72:	
30	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 504 base pairs (B) TYPE: nucleic acid	
35	(C) STRANDEDNESS: double	
,,	(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 72:	
40	GCAGGGCGA GGGGYTGGGG ACCGCGGGGC GGACGGGAGC GAGTATGTCC GCTCTGACTC	60
	GGCTGGCGTC TTTCGCTCGC GTTGGAGGCC GCCTTTTCAG AAGCGGCTGC GCACGGACTG	120
	CTGGAGATGG TGGAGTCCGT CATGCCGGTG GTGGTGTGCA CATTGAGCCC CGGTATAGAC	180
45	AGTTCCCCCA GCTGACCAGA TCCCAGGTGT TCCAGAGCGA GTTCTTCAGC GGACTCATGT	240
	GGTTCTGGAT TCTCTGGCGC TTTTGGCATG ACTCAGAAGA GGTGCTGGGT CACTTTCCGT	300
50	ATCCTGATCC TTCCCAGTGG ACAGATGAAG AATTAGGTAT CCCTCCTGAT GATGAAGACT	360
	GAAGGTGTAG ACTCAGCCTC ACTCTGTACA AGAGCCAGGT GAGAATTTCA AGGATTATCG	420
	ACTTCATATT GCACATTAAA GTTACAAATT AAAGTGGCTT GGTCAAGAAT GARAAAAAA	480
55	AAAAAAATT GGGGGGGGC CCCN	504

5	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 620 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 73:	
10	GAATTCGGCA CGAGGAGGAG GGGAGGCGGG GTAAGTTTGG TGGGAAACTC TGTAATTTCC	60
	WITTITACTT TCACAGCAAT AGTGCAGAAT CCAGAATGGA TGTCCTCTTT GTAGCCATCT	120
15	TTGCTGTGCC ACTTATCCTG GGACAAGAAT ,ATGAGGATGA AGAAAGACTG GGAGAGGATG	180
13	AATATTATCA GGTGGTCTAT TATTATACAG TCACCCCCAG TTATGATGAC TTTAGTGCAG	240
	ATTTCACCAT TGATTACTCC ATATTTGAGT CAGAGGACAG GCTGAACAGG TTGGATAAGG	300
20	ACATAACAGA AGCAATAGAG ACTACCATTA GTCTTGAAAC AGCACGTGCA GACCATCCGA	360
	AGCCTGTAAC TGTGAAACCA GTAACAACGG AACCTCAGAG TCCAGATCTG AACGATGCCG	420
	TGTCCAGTTT GCGAAGTCCT ATTCCCCTCC TCCTGTCGTG TGCCTTTGTT CAGGTGGGGA	480
25	TGTATTTCAT GTAGAAGGTG GAAGAAGGCT GCTATGACTC TTTGGATGGG AGTCTGGCAA	540
	GAGGAAATTG GAAGATAAAA TAAATAATAA GTGAAATAAA AAAAAAAA	600
30	GGGGGGCCC GGTACCCAAT	620
	· ·	
25	(2) THEODMARTON FOR SEC TO MO. 74.	
35	(2) INFORMATION FOR SEQ ID NO: 74:	
35 40	(2) INFORMATION FOR SEQ ID NO: 74: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 581 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 581 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: double	·
	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 581 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
40	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 581 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 74:	60 120
40 45	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 581 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 74: ACAAGGTGTG TGTAAAGTTT ATGTTTGTAA ACTGAATTCT ATCTTAAATC CAAAAAGAAC	120
40	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 581 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 74: ACAAGGTGTG TGTAAAGTTT ATGTTTGTAA ACTGAATTCT ATCTTAAATC CAAAAAGAAC TCGGGAGTAA TTCATTTTTG TAGCATAAAG ATCCCTAAGT TTTATTTTGA AATATCTGAT	120
40 45	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 581 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (Xi) SEQUENCE DESCRIPTION: SEQ ID NO: 74: ACAAGGTGTG TGTAAAGTTT ATGTTTGTAA ACTGAATTCT ATCTTAAATC CAAAAAGAAC TCGGGAGTAA TTCATTTTTG TAGCATAAAG ATCCCTAAGT TTTATTTTGA AATATCTGAT TTTTACACGT TAAAAAATAA CAGGGCATCG AGAGGATTCC TAGGTGACAT CCAGACTCCT	120 180
40 45	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 581 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (Xi) SEQUENCE DESCRIPTION: SEQ ID NO: 74: ACAAGGTGTG TGTAAAGTTT ATGTTTGTAA ACTGAATTCT ATCTTAAATC CAAAAAGAAC TCGGGAGTAA TTCATTTTTG TAGCATAAAG ATCCCTAAGT TTTATTTTGA AATATCTGAT TTTTACACGT TAAAAAATAA CAGGGCATCG AGAGGATTCC TAGGTGACAT CCAGACTCCT TTAGCTTTGT GTGTGTGGCA CCGGTTAGTC TGCTTCTCTC TCCTTTCTTG CACTGCTTCA	120 180 240
40 45 50	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 581 base pairs (B) TYPE: nucleic acid (C) STRANDEINESS: double (D) TOPOLOGY: linear (Xi) SEQUENCE DESCRIPTION: SEQ ID NO: 74: ACAAGGTGTG TGTAAAGTTT ATGITTGTAA ACTGAATTCT ATCTTAAATC CAAAAAGAAC TCGGGAGTAA TTCATTTTTG TAGCATAAAG ATCCCTAAGT TTTATTTTGA AATATCTGAT TTTTACACGT TAAAAAATAA CAGGGCATCG AGAGGATTCC TAGGTGACAT CCAGACTCCT TTAGCTTTGT GTGTGTGGCA CCGGTTAGTC TGCTTCTCTC TCCTTTCTTG CACTGCTTCA CACAGCCATG CCCTGCCAGC CCGGGCAGGT GCCTTCCTGT CAATGTACAT TTGGGCTTCT	120 180 240 300

	CCGSTAAAGC CATAAACTCC TTAAGGACAG GTAGCATTCT TAGTATCTTC GTTCTTCTCA	540
	ATGACCAGTA GACCATTAAA CATGTAGCAA ACAAATGTGA A	581
·5	, '	
	(2) INFORMATION FOR SEQ ID NO: 75:	
10	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1843 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double	
15	(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 75:	
	AAACCCAACN CCCTCCGGTC CCCNAAAGAA AGCCCAGCCC AAATCCCAAG CCGGCAGTGA	60
20	GCCCGCGAAC AAGGCCCTCA AGACGCCCAG NCGAACAAGC AGCCCCCAGG AGGCCCCGCA	120
	AGAGAACTCC CTGGCGGCCC AAGCGGGCAG CTTCTGTGCG GCAGAACTCA GCCACCGAGA	180
25	GCGCAGACAG CATCGAGATT TATGTCCCGG AGNCCCAGAC CAGGCTCTGA GACCATGCAG	240
	GAGGAAAGAA ACGATTTTAA ATCATTAAAA ACACAAAAAC TAAGTGCGAA CGGAACAGAG	300
	TTTTCTCAAC CTTTGCTATG GTTATTCTGT CTAGAGACCC TGAGCCAACT TTCAAATTGA	360
30	CGCATACAAG GGCTCACAAT TIGGCTTTTT TGGGTCCCTC CCAGCTTTAG GTTATGAAGA	420
	TTTTACTCAC AAAAAAAATC AACAAAAATC ACGAAACTAG AAAACTTTTT TTTTCCTCTT	480
35	GCTGGCCGTG GTGGACTAGA TAGATGGACG TCGGCAACTC CCGGCCCAGC CTCCATACTG	540
<i>55</i>	CGGTCTTTTT ACTCGTTCTA TCTGATGAGA ACTCACACTA GCTTGTTTAC AAGATGACGA	600
	CAGTCCAAGG GCAGCCTTGG GCACCTGCCA TGTCCCTCCT TTCCCCAGCT ATCCCCGCTC	660
40	TGACCITGAT TITCATTCTT ATGITTTTCT CTTTTCCCTT CAGAGCTCAC ACAGTGGTCA	720
	CCATTGTGGC AAGCGGCTTT CTGGGTCTCA GCCCTCTCTG CGGTTGAGGG CCCAGAGGAC	780
	AGAGAGATGG ACATGCGTCC CCTCCCTCCC CCCGCCAAGT GCTCACACAC AACCTCACGC	840
45	GCACACACA ACACGCAGAT GGAGGCGCCT CACTGGGAGG TGCCCCGCCA GCCCTGGGCA	900
	GTGTCAGGCA GGACTCACTC ACCGCTGAGC AGATGAGAGA AGTTTTAGTC TTGGCGGGTG	960
50	GAAATGAGAC GAAGCCACAG TTATCACACT CCAGACTCCT GCCCTTTTAT TTTCTCCAGC	
	CCCTTCTTCC TTCAGCAAAA TCTAGGACTC CCGAGTGGCT TCCAGGGGGC CGTCAGTCCT	
	CAGCCGCGCC TGTGTCCGGT GCCCGAGGGG CGGGCGGCGG TGTCTGTATG TATGTGTACA	
55		
	TATGCACATA GACCTTAGAG TGTATAGTTA ACAAACGCCC ATCTGCTCAC CCATGCCCAC	
60	CCAGCGCCGC CGCCGCTGGC TCTCGGGGCA CCTGGCAGGA GGCGGGTGTG TGAATAGCAT	
. / . /	ALACTETER ATTENDED TO THE PROPERTY OF THE PROPERTY OF THE PROPERTY AND ASSESSMENT OF THE PROPERTY OF THE PROPE	1200

	GTGTGTAAGC AGCCCTTTTT TTTTTTGGTC TCCACCCCCC TCCCCCGGCC CCGCACTCCT	1380
5	AAGGCCCCAT CTGCCCCAGCC TCTGAGTTTT CTGTTCTATT TTTTTTTTTAA CCCCCAATTAT	1440
	CCTTCTCTCT CTCCTGCCCC CGCATCCCAC TCCCAGGGTG TCACGAGCCC TGAGCTGCAA	1500
	TGGCCCGGGC CTGCAGGGCG GGGTAGGGGA GGGCARGGCT SAGCCCCGAA GCCAGCTCAG	. 1560
10	TACCTGAGGG GCTGCTCTAT GCTGTGTATG CGCCTCTCTG GCATCCGAGA CATCCTCTTG	1620
	GTGGCGCTTG CTNGCAGGGG ACCCCCCCCC CGTCCCCAGG TGAACCAAGG GTCTGCTCCG	1680
15	GGGCCCATTT CCAGCTTGGC CGCCGTCTGT GACCTTGGGC AAGTCACTTG ACCTCTGTGT	1740
••	GCCTCAACTT CCTCCTCTGT AAAACGGGGA CAGTCCCTGC CCCTCCCTAC CTCACAGGCA	1800
	TGTTGTGAGA ATAAATGAGG TAACGTGTAA AAAAAAAAAA	1843
20		
	(2) INFORMATION FOR SEQ ID NO: 76:	
25	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 1441 base pairs (B) TYPE: nucleic acid	•
	(C) STRANDEDNESS: double (D) TOPOLOGY: linear	
30	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 76:	
	TOGACCCACG CGTCCGGCTC CCCGAGCCCT GCCAACCATG GTGAACTTGG GTCTGTCCCG	60
35	GGTGGACGAC GCCGTGGCTG CCAAGCACCC GGGACTCGGG GAGTATGCCG CATGCCAGTC	120
	ACACGCCTTC ATGAAGGGCG TTTTCACCTT CGTCACAGGC ACCGGCATGG CCTTTGGCTT	180
40	GCAGATGTTC ATTCAGAGGA AGTTTCCATA CCCTTTGCAG TGGAGCCTCC TAGTGGCCGT	240
70	GGTTGCAGGC TCTGTGGTCA GCTACGGGGT GACGAGAGTG GAGTCGGAGA AATGCAACAA	300
	CCTCTGGCTC TTCCTGGAGA CCGGGCAGCT CCCCAAAGAC AGGAGCACAG ATCAGAGAAG	360
45	CTAGGAGAGC TCCAGCAGGG GCACAGAGGA TTGGGGGCAG GAGGAGTCTG GAACACAGCC	420
	TTCATGCCCC CTGACCCCAG GCCGACCCTC CCCACACCCT AGGGTACCCC AGTCGTATCC	480
50	TCTGTCCGCA TGTKTGGCCA GGCCTGACAA ACACCTGCAG ATGGCTGCTG CCCCAACCTG	540
50	GGACCIGCCC AGRAGGITGG AGCAGAAAGG GCTCTCCCTG GGGTGGTGTT TCTCCTCTAG	600
	GGTATTGGGA TGCATGTTCT GCACTGCCAG CAGAGAGGGT GTGTCTGGGG GCCACCACCT	660
55	ATGGGACACG GGGTCGAAGG GGCCTGTACA CTCTGTCATT TCCTTTCTAG CCCCTGCATC	720
	TCCAACAAGT CCAAGGTGAC AGCTGGTGCT AGGGGCGTGG GGTTAATAAA TGGCTTATCC	7.80
60	TTCTCTCCAC CCAAGTTTCC ACCTGACCAG GTGAAAAACA AATCAGAAGG GTAAGATGAT	840

	GACAGGTCAC	ATGAAACCTT	TATTACCCTA	CAGTTGATAT	ATGAGGATCA	CATGCAAGTT	900
	ACATACTGAG	GATGTACAGG	GAAGTTCCCA	GCGCTGAACC	CCAGAATTAG	ACGTTCGCAT	960
5	CAGCCCCGTA	GGCCACGTGG	ACACCACCAC	AGCCTCTCTG	TATGGGGGTC	TGCCTCTGTA	1020
	GCACTTGGCA	TGTAGGGGCA	GAGCAAAAGG	GCCANGCTG	GCCAGAGCCT	GGCTGCTGGG	1080
10	NAGARGAGGG	ACTIGIGGGS	CACGCCACNT	GCCTATCATT	CCCCAYTCAT	CTATTAGCCA	1140
•	AAGTCACTCC	CCAGAGGCAG	AGCTAGCCCG	TTGTAGCCGT	GICIGIGIGG	AGGGAAAGCT	1200
	TCTGAGTGGG	CAAGCCTACA	CACAGCCCCG	AGCCCCAAGA	GGAGGAAGAG	GTGGAGACCA	1260
15	GACGGAACCT	CCACAAGTCC	ATCATGGTTA	CAGCIGGCTT	CCCCGCAGCA	CCGAAGACCC	1320
	ACAGCATNGG	CCCTGCTGCC	CCCGACCCAG	CTCAGCTGCC	ANGCCTCACC	TTGCCAGGAA	1380
20	TTGAAAGAAA	GTTATTGAGT	ACTAATTGGC	CTCAGAGTNA	CAGGAAGCTC	AAGTTAAAGT	1440
	G					ı	1441

30

(2) INFORMATION FOR SEQ ID NO: 77:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 910 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 77:

35 GGCAGAGCTG GCCTTCGACT CGCTATGTCC ACTAACAATA TGTCGGACCC ACGGAGGCCG 60 AACAAAGTGC TGAGGTGAGG ACCCCAGCGT CGTGGGCACG GGTTCGGGTT GTGGGTGTGG 120 40 ATCGGGGCCC TGGGAAGCGC CTGTCTATCC CGGGGGCAGG ACCTGAGCGC CCCTGACCCT 180 CGAGCCTGTC GCAGGTACAA GCCCCCGCCG AGCGAATGTA ACCCGGCCTT GGACGACCCG 240 ACGCCGGACT ACATGAACCT GCTGGGCATG ATCTTCAGCA TGTGCGGCCT CATGCTTAAG 300 45 CTGAAGTGGT GTGCTTGGGT CGCTGTCTAC TGCTCCTTCA TCAGCTTTGC CAACTCTCGG 360 AGCTCGGAGG ACACGAAGCA AATGATGAGT AGCTTCATGT GAGACTTGCC CTACAGAACA 420 50 AGTGACTCTT GAGTAAGGGG TGGGGGGACC CCAGCCTGGC CATCCTAGAC TGACACCTCT 480 CTCCTGTCTT CATGCTGTCC ATCTCTGCCG TGGTGATGTC CTATCTGCAG AATCCTCAGC 540 CCATGACGCC CCCATGGTGA TACCAGCCTA GAAGGGTCAC ATTTTGGACC CTGTCTATCC 600 55 ACTAGGCCTG GGCTTTGGCT GCTAAACCTG CTGCCTTCAG CTGCCATCCT GGACTTCCCT 660 GAATGAGGCC GTCTCGGTGC CCCCAGCTGG ATAGAGGGAA CCTGGCCCTT TCCTAGGGAA 720 60 CACCCTAGGC TTACCCCTCC TGCCTCCCTT CCCCTGCCTG CTGCTGGGGG AGATGCTGTC 780

	CATGITICTA GGGGTATICA TITGCTITCT CGTTGAAACC TGTTGTTAAT AAAGITTTTC	· 840
5	ACTCTGAAAA AAAAAAAAA AAAAAAAAAC TYGRGGGGG GCCCGGAACC CAATTCSCCG	900
	GATAGTGAGT	910
10		
10	(2) INFORMATION FOR SEQ ID NO: 78:	
15	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 2776 base pairs (B) TYPE: nucleic acid (C) STRANDELNESS: double (D) TOPOLOGY: linear	•
20	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 78:	
	TCGACCCACG CGTCCGGGCG GGCAGTGATG GCGGCTGGTG ATGGGGACGT GAAGCTAGGC	60
	ACCCTGGGGA GTGGCAGCGA GAGCAGCAAC GACGGCGGCA GCGAGAGTCC AGGCGACGCG	120
25	GGAGCGGCAG CGRAAGGGGG AGGCTGGGCG GCGGCGGCT TGGCGCTTCT GACGGGGGC	180
	GGGGAAATGC TGCTGAACGT GGCGCTGGTG GCTCTGGTGC TGCTGGGGGC CTACCGGCTG	240
30	TEGETECECT GEGGEGGGG GGGTCTGGGG GCCGGGGCCG GGGCGGGGGA GGAGAGCCCC	300
	GCCACCTCTC TGCCTCGCAT GAAGAAGCGG GACTTCAGCT TGGAGCAGCT GCGCCAGTAC	360
	GACGGCTCCC GCAACCCGCG CATCCTGCTC GCGGTCAATG GGAAAGTCTT CGACGTGACC	420
35	AAAGGCAGCA AGTTCTACGG CCCGGCGGGT CCATATGGAA TATTTGCTGG TAGGGATGCC	480
	TCCAGAGGAC TGGCCACATT TTGCCTAGAT AAAGATGCAC TTAGAGATGA ATATGATGAT	540
40	CTCTCAGATT TGAATGCAGT ACAAATGGAG AGTGTTCGAG AATGGGAAAT GCAGTTTAAA	600
	GAAAAATATG ATTATGTAGG CAGACTCCTA AAACCAGGAG AAGAACCATC AGAATATACA	660
	GATGAAGAAG ATACCAAGGA TCACAATAAA CAGGATTGAA CTTTGTAAAC AACCAAAGTC	720
45	AGGGCCTTC AGAACTGCAA TICTTACTCC CTTTCACAGA CTGTCCGGAG TCTTTGGGTT	780
	TGATTCACCT GCTGCGAAAA ACATTCAACA AATTGTGTAC AAGATAAATT AATCTCACTA	840
50	TGAAGATTTG AATAACTAGA CATTATTTAT GCTGCCAAAC TCATTTGTTG CAGTTGTTTG	900
	TAATGICTAG TGGGGCTTCA TCATCCTGAA AAGAAGGAGA CAGGGATTIT TITAAAGAGC	960
e e	AAGAAAGTCA CAATATTACT TCTTTCCTTC CTTTTTTCCT TCTTTCCTTT CTTCTT	1020
55	TTTCTTTCTT TTTAAAATAT ATTGAAGACA ACCAGATATG TATTTGCTAC TCAAGTGTAC	1080
	AGATCTCCTC AAGAAACATC AAGGGACTCC TGTGTCACAT ACTGTGTTT TATTTTAACA	1140
60	TGGGTGAGGG AGGCGACCTG ATCAGGGGGAG GTGGGGGGTAC ACATCAATTT GAGTTGTTCA	1200

	GGCTACTGAA ACATTAAAAT GTGAATTCCC AAACTTTTCT TTTTGGCTTT GTCAGGGAAA	1260
	AGAAAAATAT CTTTATAAAG AAATCTTTGG AAATTAGGAG AAGGAATTC AGGTGGTTT	1320
5	AAGTCAGAGC TAGTTCCCCA ACAGAAAGAT CATTTGAAAC CAGTTTTTAT CCCTTCTCTT	1380
	TCCTTCCCTT TCCCTAAATC AAATCAATAT TAATTGTGCC TTATTTCACT TAACATAGAC	1440
10	TTGAATTATT TTTAGGGAAA GCCCCTATAA TGAATTCAGA AATCACTACA AGCAGCATTA	1500
10	AGACTGAAGT TGGAATATTC TGTTGACCAT AAAACCTTGA TATCATTCTG TGTATATAGA	1560
	ATGTAAAAGG AATATTACAG TGTTAACTGC CATATATGTA ATATACACAA ACTCAATTAG	1620
15	CATTGTAATG GCCAAATGCA TTCCCCCATG CTTTTCTGTT TTCAAAAAAA TTGAAAAACA	1680
	AATCAACTCT TATCCCCAAC AGCTGCCTAA TITTAGGAGT CTGACCCTCC ACATCTCACT	1740
20	GGTGTGGGTG CATGGGGCTG TGGAGTGGGT GTCAGTATGG ATGTGTCTGA ATGTGTGAGG	1800
20	CCTTGGAAGG GACTCTTTCT GCAGATACTG TAAATACAAG TACCATTTTA ATAAAGCATG	1860
	TACAATAAAC CAAAATAAGC TTGAGTTGGA CTTTATATAC AGAACTGTAA GCCAGTGCAT	1920
25	TATGATACAG TIGTAAGATT GIGCATTIGA TICAAGATAA GGAAAAATCT TGGAAATGAA	1980
	AAGCAGGCAC KGGTTAACCA AGTTGTACAC ATTGTACCAC ATTCAGCATA ACTTTAGGAA	2040
30	GAAATTCCAC TTTGTGAACA TTCTCCAGAA ATCCAAGATT ATTCAGGTAA GAATTGGTAT	2100
50	ATTAAATGTA CATCTTTTTA CTTTCTATTT TGATGCCAAC TGATTATACT AGACAATTAG	2160
	CACTCCAGGT GGTTATTGAA CACAAAACAG TAAAAGAATA TTGCACTGAT AGATACTAAA	2220
35	TTATTATTTT ATTAGGTTGA AAAAGCCCTT ACTAAAAGCC CCTCATATAT CAATTACTTT	2280
	ATTTCATTAT GACTACTTAG GTTCCGGGCT GGGGACAAGT TCACTTAAAA AGGCAATGTT	2340
40	ATTTAACAGG TCACCAGTTA AGACTTCTGC TTTGTAGATA CATGCAGAAG CCATCAAACA	2400
	AGGGGGRGCT TITAACTGCA ACAATAAGCT AAAGTATGTA AAATACTACA TICTATTCAG	2460
	TCTTGGAGTG TTTTGTAGAA AGITATCTTC AGCCAAATCT TTGCTGAAGA CTGGTTGTGG	2520
45	AGIGITIGGIA AATGCTTTIGI GITTITATIGI AAAATATITT CTAAACAAAA AATGTTAAAA	2580
	GTACATGTCC TCTGTAGTAA ACTGATATCT ATATATATGA ATCATTCAAG CCTAAAGTCT	2640
50	AGTAATAAAC TGTACTTGTG AATAGAGAAA CCCTAAATAT TCATGCAGWA AAAATTATGC	2700
	GGTCTGTTAA GAAAAATGAG TAATTTGTGT TTTGGACTTG AAATAAACAG TGTTCTGTAG	2760
	ATAATTCCTC AACTTC	2776

⁽²⁾ INFORMATION FOR SEQ ID NO: 79:

(A) LENGTH: 1525 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 79:

					•		
	CCGCTGCTGA	TAACTATGGC	ATCCCCCGGG	CCTGCAGGAA	TTCGGCACGG	AGCTACGGCG	60
10	CCGCCTGGCT	CCTGCTGNCA	CCTGCAGGCT	CCTCCCCCCT	GGAGCCCACC	CAAGACATCA	120
	GCATCAGCGA	CCAGCTGGGG	GCCAGGACG	TGCCCGTGTT	CCGGAACCTG	TCCCTGCTGG	180
15	TOGTGGGTGT	CGCCGCCGTG	TTCTCACTGC	TATTCCACCT	GGGCACCCGG	GAGAGGCGCC	240
13	GGCCGCATGC	GGASGAGCCA	GGCGAGCACA	CCCCCTGIT	GCCCCIGCC	ACGGCCCAGC	300.
	CCCTGCTGCT	CTGGAAGCAC	TGGCTCCGGG	AGCSGGCTTT	CTACCAGGTG	GGCATACTGT	360
20	ACATGACCAC	CAGGCTCATC	GTGAACCTGT	CCCAGACCTA	CATGGCCATG	TACCTCACCT	420
	ACTCGCTCCA	CCTGCCCAAG	AAGTTCATCG	CGACCATTCC	CCTGGTGATG	TACCTCAGCG	480
25	GCTTCTTGTC	CICCIICCIC	ATGAAGCCCA	TCAACAAGTG	CATTGGGAGG	AACATGACCT	540
23	ACTTCTCAGG	CCTCCTGGTG	ATCCTGGCCT	TTGCCGCCTG	GCTGCCCCTG	GCGGAGGGAC	600'
	TGGGTGTGGC	CGTGTACGCA	GCGGCTGTGC	TGCTGGGTGC	TEGETETECC	ACCATCCTCG	660
30	TCACCTCGCT	GGCCATGACG	CCCGACCTCA	TCGGTCCCCA	CACGAACAGC	GGACTKTCGT	720
	GTACGGCTCC	ATGAGCTTCT	TGGATAAGGT	GGCCAATGGG	CTGGCAGTCA	TGGCCATCCA	780
35	GAGCCTGCAC	CCTTGCCCCT	CAGAGCTCTG	CTGCAGGGCC	TGCGTGAGCT	TTTACCACTG	840
55	GGCGATGGTG	GCTGTGACGG	CCCCCCTCCC	CGTGGCCGCT	GCCCTGTGTC	TCTGTAGCCT	900
	CCTGCTGTGG	CCGACCCGCC	TGCGACGCTG	GGACCGTGAT	GCCCGGCCCT	GACTCCTGAC	960
40	AGCCTCCTGC A	ACCTGTGCAA	GGGAACTGTG	GGGACGCACG	AGGATGCCCC	CCARGGCCTT	1020
	GGGGAAAAGC	CCCCACTGCC	CCTCACTCTT	CTCTGGACCC	CCACCCTCCA	TCCTCACCCA	1080
45	GCTCCCGGGG (GTGGGGTCGG	GTGAGGGCAG	CAGGGATGCC	CGCCAGGGAC	TTGCAAGGAC	1140
73	CCCCTGGGTT	PTGAGGGTGT	CCCATTCTCA	ACTCTAATCC	ATCCCAGCCC	TCTGGAGGAT	1200
	TTGGGGTGCC (CCTCTCGGCA	GGGAACAGGA	AGTAGGAATC	CCAGAAGGGT	CTGGGGGAAC	1260
50	CCTAACCCTG	AGCTCAGTCC	AGTTCACCCC	TCACCTCCAG	CCTCGGGGTC	TCCAGACACT	1320
	GCCAGGGCCC (CCTCAGGACG	GCTGGAGCCT	GGAGGAGACA	GCCACGGGGT	GGTGGGCTGG	1380
55	GCCTGGACCC (CACCGTGGTG	GGCAGCAGGG	CTGCCCGGCA	GCTTGGTGG	ACTCTGCTGG	1440
JJ	CAGCAAATAA A	AGAGATGACG (GCAAAAAAA	AAAAAAAAA		АААААААА	1500
	AAAAAAAAA	AAACCCACCG	recec				15 25

(2) INFORMATION FOR SEQ ID NO: 80:

5	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1563 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double	
10	(D) TOPOLOGY: linear	•
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 80:	
	AATTCGGCAC GAGNCAGAAA CCTGCGGAAA ATGGTAGCGA TGGCGGCTGG GCCGAGTGGG	60
15	TENCHEGIEC CECCETTEE CTACEGITE TIGTTECCEA CTETECTICA ACCESTETCT	120
	GCTTTTGGGG CAGAGTTTTC ATCGGAGGCA TGCAGAGAGT TAGGCTTTTC TAGCAACTTG	180
20	CTTTGCAGCT CTTGTGATCT TCTCGGACAG TTCAACCTGC TTCAGCTGGA TCCTGATTGC	240
20	AGAGGATGCT GTCAGGAGGA AGCACAATTT GAAACCAAAA AGCTGTATGC AGGAGCTATT	300
	CTTGAAGTTT GTGGATGAAA ATTGGGAAGG TTCCCTCAAG TCCAAGCTTT TGTTAGGAGT	360
25	GATAAACCCA AACTGTTCAG AGGACTGCAA ATCAAGTATG TCCGTGGTTC AGACCCTGTA	420
	TTAAAGCTTT TGGACGACAA TGGGAACATT GCTGAAGAAC TGAGCATTCT CAAATGGAAC	480
30	ACAGACAGTG TAGAAGAATT CCTGAGTGAA AAGTTGGAAC GCATATAAAT CTTGCTTAAA	540
50	TITTGTCCTA TCCTTTTGTT ACCTTATCAA ATGAAATATT ACAGCACCTA GAAAATAATT	600
	TAGTTTTGCT TGCTTCCATT GATCAGTCTT TTACTTGAGG CATTAAATAT CTAATTAAAT	660
35	CGTGAAATGG CAGTATAGTC CATGATATCT AAGGAGTTGG CAAGCTTAAC AAAACCCATT	720
	TTTTATAAAT GTCCATCCTC CTGCATTTGT TGATACCACT AACAAAATGC TTTGTAACAG	780
40	ACTTGCGGTT AATTATGCAA ATGATAGTTT GTGATAATTG GTCCAGTTTT ACGAACAACA	840
70	GATTICTAAA TTAGAGAGGT TAACAAGACA GATGATTACT ATGCCTCATG TGCTGTGTGC	900
	TCTTTGAAAG GAATGACAGC AGACTACAAA GCAAATAAGA TATACTGAGC CTCAACAGAT	960
45	TGCCTGCTCC TCAGAGTCTC TCCTATTTTT GTATTACCCA GCTTTCTTTT TAATACAAAT	1020
	GTTATTTATA GITTACAATG AATGCACTGC ATAAAAACTT TGTAGCTTCA TTATTGTAAA	1080
50	ACATATTCAA GATCCTACAG TAAGAGTGAA ACATTCACAA AGATTTGCGT TAATGAAGAC	1140
30	TACACAGAAA ACCTITCTAG GGATTTGTGT GGATCAGATA CATACTTGGC AAATTITTGA	1200
	GITTTACATT CTTACAGAAA AGTCCATTTA AAAGTGATCA TTTGTAAGAC ÇAAAATATAA	1260
55	ATAAAAAGTT TCAAAAATCT ATCTGAATTT GGAATTCTTC TGGTTTGTTC TTTCATGTTT	1320
	AAAAATGATG TTTTTCAATG CATTTTTTTC ATGTAAGCCC TTTTTTTAGC CAAAATGTAA	1380
	AND CONTRACT AND CONTRACTAR AND ACCOUNTAGE OF THE CONTRACTOR CONTRACTOR CONTRACTOR	1440

	GICIGATITY	ATTITICAAA	GITTITCAT	TTATGAACAC	ATTTTCATTG	GTATATTATT	1500
	TAAGGAATAT	CTCTTGATAT	AGAATTTTTA	TATTAAAAAT	GATTTTTCTT.	TGCTTAAAAA	1560
5	AAA						1563

10 (2) INFORMATION FOR SEQ ID NO: 81:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1020 base pairs.

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear '

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 81:

20	TGCACGCTGG	CCATGTGGGN	GTTGGGCCAC	TGCGACCCCC	GGCGCTGCAC	GGCCGCAAG	60
	CIGGCCCCCC	TGGGGCTGGT	GCGCTGCCTG	CCCCTCCCCC	ACAGATTCGG	CGGTCTGGTG	120
25	CTGAGCCCCG	TGGGCAAGCA	GTACGCGTCC	CCCGCAGACA	GACAGCTGGT	GCCCACTCT	180
	GGGTCGCCG	TCATCGACTG	CTCCTGGGCC	AGGCTGGACG	AGACACCGTT	TGGGAAGATG	240
	CGAGGGAGCC	ACTTGCGCCT	GTTGCCCTAC	CIGGIGGCCG	CCAACCCCGT	GAACTATGGC	300
30	CGGCCCTACA	GACTTTCCTG	CGTGGAAGCG	TTTGCTGCCA	CCTTCTGCAT	CGTAGGCTTT	360
	CCAGACCTTG	CTGTCATTTT	GCTGCGGAAG	TTTAAATGGG	GCAAGGGCTT	CTTGGACCTG	420
35	AACCGCCAGC	TCCTGGACAA	GTACGCGGCC	TGCGGCAGCC	CGGAGGAGGT	GCTGCAGGCG	480
33	GAGCAGGAGT	TCTTGGCCAA	TGCCAAGGAG	AGCCCCCAGG	AGGAGGAGAT	CGATCCCTTC	540
	GATGTGGATT	CAGGGAGAGA	GTTTGGAAAC	CCCAACAGGC	CTGTGGCCAG	CACCCGGCTG	600
40	CCCTCGGACA	CTGATGACAG	TGATGCGTCT	GAGGACCCAG	GCCTKGCGC	CGAGCGCGGA	660
	GGAGCCAGCA	GCAGCTGCTG	TGAAGAGGAG	CAGACGCAGG	GACGGGGGC	TGAGGCCAGG	720
45	GCCCCGGCTG	AGGTTTGGAA	AGGAATCAAG	AAACGGCAGA	GAGACTGAGG	GTTGCAGACA	780
43	CATATATTT	TGAGGCTGGG	TGACGAGAAA	ATCTAGAGAC	ATGAGGGACA	TAAATGGGCC	840
	TGGCAGCCTC	GGCTCTTTGC	GGCTGCTGGC	AGGACTGAGC	TGTCCGGGTT	CTCCCCACAC	900
50	TTCCAGCACA	GCTGTGCTCT	GTGTCCTGCC	TCGGCGCTCT	CGCAAATGAA	GCTGCAGGCC	960
	AAGAAAAAA	АААААААА	АААААААА	АААААААА	AAAAAAAAG	GGGGGGGGC	1020

55

60

(2) INFORMATION FOR SEQ ID NO: 82:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 770 base pairs

	·(B)	TYPE: nucleic acid	
	(C)	STRANDEDNESS: double	
	(D)	TOPOLOGY: linear	
(xi)	SEQUE	NCE DESCRIPTION: SEQ ID NO:	82:

TOGACCCACG CGTCCGGGCC GCCGTAGCGC GTCTTGGGTC TCCCGGCTGC CGCTGCTGCC 60 GCCGCCGCCT CGGGTCGTGG AGCCAGGAGC GACGTCACCG CCATGGCAGG CATCAAAGCT 10 TIGATTAGIT TGTCCTTIGG AGGAGCAATC GGACTGATGT TTTTRATGCT TGGATGTGCC 180 CTTCCAATAT ACAACAAATA CTGGCCCCTC TTTGTTCTAT TTTTTTACAT CCTTTCACCT 240 ATTCCATACT GCATAGCAAG AAGATTAGTG GATGATACAG ATGCTATGAG TAACGCTTGT 300 AAGGAACTTG CCATCTTTCT TACAACGGGC ATTGTCGTGT CAGCTTTTGG ACTCCCTATT 360 GTATTTGCCA GAGCACATCT GATTGAGTGG GGAGCTTGTG CACTTGTTCT CACAGGAAAC 420 20 ACAGTCATCT TTGCAACTAT ACTAGGCTTT TTCTTGGTCT TTGGAAGCAA TGACGACTTC 480 AGCTGGCAGC AGIGGTGAAA AGAAATTACT GAACTATTGT CAAATGGACT TCCTGTCATT 540 25 TGTTGGCCAT TCACGCACAC AGGAGATGGG GCAGTTAATG CTGAATGGTA TAGCAAGCCT 600 CTTGGGGGTA TTTTAGGTGC TCCCTTCTCA CTTTTATTGT AAGCATACTA TTTTCACAGA 660 720 30 GGGGGGCCC GTWCCCATTC SCCCYATATG AATTCCNTTT TTACAATCCC 770

35

40

(2) INFORMATION FOR SEQ ID NO: 83:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 481 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 83:

45 GAATTCGGCA CGAGCATAGT GTTAACCACT AGAATTCACT GCCCTTCCTA TCCAAAAATG 60 ACACTACTGA TCATTTTTCT TCCTTTTSCT TTTACAACAT TMACAAATTC AGGTGGCTCT 120 50 TTCCCAGTAC GGTAGGCTGA TTCGTATGGA TGCACCACGG TTGGTGACTC CCCCCACCCC 180 ACAGAGTITC TGGCGTTCAT TCGGTTGAAC CCAAGGCCAG CAAGGGCTGA CTGGGAACAA 240 ACCGAACACT AGGCCGTGAA CCAATCGTCT CTCCGTGCCC GGGAGCGAMC CCGGGGGCCT 300 55 TTCACTCTCC CAAGGACTCC ANGGGGGGC CGGGTACCCA ATTCCGCCCC TATAGTGAAT 360 CCGTNATTAC AATTCCACNT GGGCCGTCCN TTTTTACAAA CGTTCCGTTG AACTGGGAAA 420 60 AACCCCTTGG CGGTTTACCC CAACTTTAAT CCGCCTTTGC AAGCACATCC CCCCCCTTTT 480

	С	481
5		•
	(2) INFORMATION FOR SEQ ID NO: 84:	
10	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 644 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
15	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 84:	
	GCTGGGATAG AGCATGAAAG GAGAACTGCT CCCTTTTCTG TTTCTCACAG TTTGGTTATG	60
20	GCTTTATAAA CTTKTATTTG GTGAAAGCCC CAGATACCCA AATGTCATTG GCAAAACTTA	120
20	TTTTTTTTC TGGACAGATC AGATTTCTAG AGAGAGCAGA TTTCTAGAGA GATTAGCATT	180
	CATAGTAAGT GAAAATTGTC TAATTTTTTT AATCCATGCT ATTACTGGGC AGTAGGTCTA	240
25	ATTITITIG ACAAAAAATA GATCTATITI CCTTATATAT TGATTTAGAA TCTTAAGITA	300
-	GAATTITATA GAAGAAATGI CIGAGCAGII CIATGIATGG AGGAGCAATI CAGCIITICA	360
30	GCAGCAACTT TATCTTTTGC CACTAGAGGG AGATCTGTGG TTGCTTTCTC CTTTGGAGAA	420
30	TAGCTGCTTT GCTTTTATTT TTAATTTCTA AGGTTGGAAT AGAACTTATT CTCAAAATTC	480
	CTTTAGTGTT ATTAAATATT TTCATTTATT AGTCAAAGGT AAGTTAATTA AGCTTGTTTA	540
35	ATGATGCCAA TCTTATGCTT TTCTGTAATC TTCAATTTTT AATAAATGTG AGTTAGATAC	600
	TAAGTGAAAA AAAAAAAAA AAAAAAAAA AAAAAAAAA	644
40		
10	(2) INFORMATION FOR SEQ ID NO: 85:	
45	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1351 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear 	
50	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 85:	
	GCCACGAGTG CGCCASGCGTG GGGCTCTCTC CTTGTCAGTC GGCGCCGCGT GCGGGCTGGT	60
E E	GGCTCTGTGG CAGCGGCGC GGCAGGACTC CGGCACTATG AGCGGCTTCA GCACCGAGGA	120
55	GCGCGCCGCG CCNTTCTCCC TGGAGTACCG AGTCTTCCTC AAAAATGAGA AAGGACAATA	180
	TATATCTCCA TITCATGATA TTCCAATITA TGCAGATAAG GATGTGTTTC ACATGGTAGT	240
60	TGAAGTACCA CGCTGGTCTA ATGCAAAAAT GGAGATTGCT ACAAAGGACC CTTTAAACCC	300

	TATTAAACAA GATGIGAAAA AAGGAAAACT TCGCTATGTT GCGAATTTGT TCCCGTATAA	360
·5	AGGATATATC TGGAACTATG GTGCCATCCC TCAGACTTGG GAAGACCCAG GGCACAATGA	420
	TAAACATACT GGCTGTTGTG GTGACAATGA CCCAATTGAT GTGTGTGAAA TTGGAAGCAA	480
	GGTATGTGCA AGAGGTGAAA TAATTGGCGT GAAAGTTCTA GGCATATTGG CTATGATTGA	540
10	CGAAGGGGAA ACCGACTGGA AAGTCATTGC CATTAATGTG GATGATCCTG ATGCAGCCAA	600
	TTATAATGAT ATCAATGATG TCAAACGGCT GAAACCTGGC TACTTAGAAG CTACTGTGGA	660
15	CTGGTTTAGA AGGTATAAGG TTCCTGATGG AAAACCAGAA AATGAGTTTG CGTTTAATGC	720
15	AGAATTTAAA GATAAGGACT TTGCCATTGA TATTATTAAA AGCACTCATG ACCATTGGAA	780
	AGCATTAGIG ACTAAGAAAA CGAATGGAAA AGGAATCAGT IGCATGAATA CAACTTIGIC	840
20	TGAGAGCCCC TTCAAGTGTG ATCCTGATGC TGCCAGAGCC ATTGTGGATG CTTTACCACC	900
	ACCCTGTGAA TCTGCCTGCA CAGTACCAAC AGACGTGGAT AAGTGGTTCC ATCACCAGAA	960
25	AAACTAATGA GATTTCTCTG GAATACAAGC TGATATTGCT ACATCGTGTT CATCTGGATG	1020
	TATTAGAAGT AAAAGTAGTA GCTTTTCAAA GCTTTAAATT TGTAGAACTC ATCTAACTAA	1080
	AGTAAATTCT GCTGTGACTA ATCCAATATA CTCAGAATGT TATCCATCTA AAGCATTTTT	1140
30	CATATCTCAA CTAAGATAAC TTTTAGCACA TGCTTAAATA TCAAAGCAGT TGTCATTTGG	1200
	AAGTCACTTG TGAATAGATG TGCAAGGGGA GCACATATTG GATGTATATG TTACCATATG	1260
35	TTAGGAAATA AAATTATTTT GCTGAAAAAA AAAAAAAAAA	1320
55	TCCCCATTIG GCCCTTTGGG GGGAGGTTTT A	1351
40	(2) INFORMATION FOR SEQ ID NO: 86:	
	(i) SEQUENCE CHARACTERISTICS:	
45	(A) LENGTH: 2527 base pairs	
45	(B) TYPE: nucleic acid	

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 86: 50

CTCTTGCTAC CTTCCCGGCG CAGAGAACCC CGGCTGCTCA GCGCGCTCCG GGGTCATGGA 60 GATCCCCGGG AGCCTGTGCA AGAAAGTCAA GCTGAGCAAT AACGCGCAGA ACTGGGGAAT 120 55 GCAGAGAGCA ACCAATGTCA CCTACCAAGC CCATCATGTC AGCAGGAACA AGAGAGGTCA 180 GGTGGTGGG ACCAGAGGTG GCTTTCGTGG TTGCACAGTT TGGCTAACAG GCTTGTCTGG 240 AGCGGGAAAG ACTACTGTGA GCATGGCCTT GGAGGAGTAC CTGGTTTGTC ATGGTATTCC 300 60

	ATGCTACACT	CTGGATGGTG	ACAATATTCC	TCAAGGICTC	CTAAAAATC	TTGGCTTTAG	360
	TCCTGAAGAC	AGAGAAGAGA	ATGTTCGACG	CATCGCAGAA	GTTGCTAAAC	TGTTTGCAGA	420
5	TGCTGGCTTA	GTGTGCATCA	CAAGTTTCAT	ATCACCTTAC	ACTCAGGATC	GCAACAATGC	480
	AAGGCAAATT	CATGAAGGTG	CAAGTTTACC	GTTTTTTGAA	GTATTTGITG	ATGCTCCTCT	540
10	GCATGTTTGT	GAACAGAGGG	ATGTCAAAGG	ACTCTACAAA	. AAAGCCCGGG	CAGGAGAAAT	60,0
10	TAAAGGTTTC	ACTOGGATCG	ATTCTGAATA	TGAAAAGCCA	GAGGCCCCTG	AGTTGGTGCT	660
	GAAAACAGAC	TCCTGTGATG	TAAATGACTG	TGTCCAGCAA	GITGIGGAAC	TTCTACAGGA	720
15	ACGGGATATT	GTACCTGTGG	ATGCATCTTA	TGAAGTAAAA	GAACTATATG	TGCCAGAAAA	780
	TAAACTTCAT	TTGGCAAAAA	CAGATGCGGA	AACATTACCA	GCACTGAAAA	TTAATAAAGT	840
20	GGATATGCAG	TGGGTGCAGG	TTTTGGCAGA	AGGTTGGGCA	ACCCCATTGA	ATGGCTTTAT	900
20	GAGAGAGAGG	GAGTACTIGC	AGTGCCTTCA	TTTTGATTGT	CTTCTGGATG	GAGGTGTCAT	960
	TAACTTGTCA	GTACCTATAG	TTCTGACTGC	GACTCATGAA	GATAAAGAGA	GGĊTGGACGG	1020
25	CTGTACAGCA	TTTGCTCTGA	TGTATGAGGG	CCGCCGTGTG	GCCATTCTTC	GCAATCCAGA	1080
	GTTTTTTGAG	CACAGGAAAG	AGGAGCGCTG	TGCCAGACAG	TGGGGAACGA	CATGCAAGAA	1140
30	CCACCCCTAT	ATTAAGATGG	TGATGGAACA	AGGAGATTGG	CTGATTGGAG	GAGATCTTCA	1200
	AGTCTTGGAT	CGAGTTTATT	GGAATGATGG	TCTTGATCAG	TATCGTCTTA	CTCCTACTGA	1260
	GCTAAAGCAG	AAATTTAAAG	ATATGAATGC	TGATGCTGTC	TTTGCATTTC	AACTACGCAA	1320
35	CCCAGTGCAC	AATGGACATG	CCCTGTTAAT	GCAGGATACC	CATAAGCAAC	TTCTAGAGAG	1380
	GGGCTACCGG	CGCCCTGTCC	TCCTCCTCCA	CCCTCTGGGT	GGCTGGACAA	AGGATGACGA	1440
40	TGITCCITIG	ATGTGGCGTA	TGAAGCAGCA	TGCTGCAGTG	TTGGAGGAAG	GAGTTCTGAA	1500
	TCCTGAGACG	ACAGTGGTGG	CCATCTTCCC	ATCTCCCATG	ATGTATGCTG	GACCAACTGA	1560
	GGTCCAGTGG	CATTGCAGAG	CACGGATGGT	TGCAGGAGCC	AACTTTTACA	TTGTTGGACG	1620
45	AGACCCTGCT	GGCATGCCTC	ATCCAGAAAC	AGGGAAGGAT	CTTTATGAGC	CAAGTCATGG	1680
	TGCCAAAGTG	CTGACGATGG	CCCCTGGTTT	AATCACTTTG	GAAATAGTTC	CCTTTCGAGT	1740
50	TGCAGCTTAC .	AACAAGAAAA	AGAAGCGTAT	GGACTACTAT	GACTCTGAAC	ACCATGAAGA	1800
	CTTTGAATTT	ATTTCAGGAA	CACGAATGCG	CAAACTTGCT	CGAGAAGGCC	AGAAACCACC	1860
	TGAAGGTTTC	ATGGCTCCCA	AGGCTTGGAC	CGTGCTGACA	GAATACTACA	AATCCTTGGA	1920
55	GAAAGCTTAG (GCTGTTAACC	CAGTCACTCC	ACCTTTGACA	CATTACTAGT	AACAAGAGGG	1980
	GACCACATAG '	PCTCTGTTGG	CATTTCTTTG	TEGTETETET	CTGGACATGC	ТТССТААААА	2040
60	CAGACCATTT !	CCTTAACTT	GCATCAGTTT	TEGTCTECCT	TATGAGTTCT	GTTTTGAACA	2100

	AGIGTAACAC ACTGATGGTT TTAATGTATC TTTTCCACTT ATTATAGTTA TATTCCTACA	2160
	ATACAATITT AAAATTGTCT TTTTATATTA TATTTATGCT TCTGTGTCAT GATTTTTTCA	2220
5	AGCTGTTATA TTAGTTGTAA CCAGTAGTAT TCACATTAAA TCTTGCTTTT TTTCCCCTTA	2280
	AAAAAAGAAA AAAATTACCA AACAATAAAC TTGGCTAGAC CTTGTTTTGA GGATTTTACA	2340
10	AGACCTTTGT AGCGATTAGA TTTTTTTTCT ACATTGAAAA TAGAAACTGC TTCCTTTCTT	2400
10	CTITICCAGTC AGCTATIGGT CTITICCAGCT GITATAATCT AAAGTATICT TATGATCTGT	2460
	GRAGGTCTG AATGAACTTC TTTACTCAAT AAAATTAATT TTTTGGCTTC TTAAAAAAA	2520
15	AAAAAA	2527
		•
20	(2) INFORMATION FOR SEQ ID NO: 87:	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 2566 base pairs (B) TYPE: nucleic acid	
25	(C) STRANDEDNESS: double	
	(D) TOPOLOGY: linear	1
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 87:	
30	CCCAAGAATT CGGCACGAGC GNGGCAWAAK TGGGATTTCT GAAACCTGTA GGCCCCAAGC	60
	CCATCAACTT GCCCAAAGAA GATTCCAAAC CTACATTTCC CTGGCCTSCT GGAAACAAGC	120
35	CATCTCTTCA CAGTGTAAAC CAAGACCATG ACTTAAAGCC ACTAGGCCGA AATCTGGGCC	180
	TACTCCTCCA ACCTCAGAAA ATGAACAGAA GCAAGCKTTT CCCAAATTGA CTGGGGTTAA	240
	AGGGAAATTT ATGTCAGCAT CACAAGATCT TGAACCCAAG CCCCTCTTCC CCAAACCCGC	300
40	CTTTGGCCAG AAGCCGCCCC TAAGTACCGA GAACTCCCAT GAAGACGAAA GCCCCATGAA	360
	GAATGTGTCT TCATCAAAAG GGTCCCCAGC TCCCCTGGGA GTCAGGTCCA AAAGCGGCCC	420
45	TTTAAAACCA GCAAGGGAAG ACTCAGAAAA TAAAGACCAT GCAGGGGAGA TTTCAAGTTT	480
45	GCCCTTTCCT GGAGTGGTTT TGAAACCTGC TGCGAGCAGG GGAGGCCCAG GTCTCTCCAA	540
	AAATGGTGAA GAAAAAAAGG AAGATAGGAA GATAGATGCT GCTAAGAACA CCTTCCAGAG	600
50	CAAAATAAAT CAGGAAGAGT TGGCCTCAGG GACTCCTCCT GCCAGGTTCC CTAAGGCCCC	660
	TTCTAAGCTG ACAGTGGGGG GGCCATGGGG CCAAAGTCAG GAAAAGGAAA AGGGAGACAA	720
55	GAATTCAGCC ACCCCGAAAC AGAAGCCATT GCCTCCCTTG TTTACCTTGG GTCCACCTCC	780
55	ACCAAAACCC AACAGACCAC CAAATGTTGA CCTGACGAAA TTCCACAAAA CCTCTTCTGG	840
	AAACAGTACT AGCAAAGGCC AGACGTCTTA CTCAACAACT TCCCTGCCAC CACCTCCACC	900
60	ATCCCATCCG GCCAGCCAAC CACCATTGCC AGCATCTCAC CCATCACAAC CACCAGTCCC	960

ATCCCATCCG GCCAGCCAAC CACCATTGCC AGCATCTCAC CCATCACAAC CACCAGTCCC

	AAGCCTACCI	CCCAGAAACA	TTAAACCTCC	GTTTGACCTA	AAAAGCCCTG	TCAATGAAGA	102
5	CAATCAAGAT	GGTGTCACGC	ACTCTGATGG	TGCTGGAAAT	CTAGATGAGG	AACAAGACAG	1086
	TGAAGGAGAA	ACATATGAAG	ACATAGAAGC	ATCCAAAGAA	AGAGAGAAGA	AAAGGGAAAA	1140
	GGAAGAAAAG	AAGAGGTTAG	AGCTGGAGAA	AAAGGAACAG	AAAGAGAAAG	AAAAGAAAGA .	1200
10	ACAAGAAATA	AAGAAGAAAT	TTAAACTAAC	AGGCCCTATT	CAAGTCATCC	ATCTTGCAAA	1260
	AGCTTGTTGT	GATGTCAAAG	GAGGAAAGAA	TGAÁCTGAGC	TTCAAGCAAG	GAGAGCAAAT	1320
15	TGAAATCATC	CGCATCACAG	ACAACCCAGA	AGGAAAATGG	TTGGGCAGAA	CAGCAAGGGG	1380
	TTCATATGGC	TATATTAAAA	CAACTGCTGT	AGAGATTGAC	TATGATTCTT	TGAAACTGAA	1440
	AAAAGACTCT	CTTGGTGCCC	CTTCAAGACC	TATTGAAGAT	GACCAAGAAG	TATATGATGA	1500
20	TGTTGCAGAG	CAGGATGATA	TTAGCAGCCA	CAGTCAGAGT	GGAAGTGGAG	GGATATTCCC	1560
	TCCACCACCA	GATGATGACA	TTTATGATGG	GATTGAAGAG	GAAGATGCTG	ATGATGGCTC	1620
25	CACACTACAG	GTTCAAGAGA	AGAGTAATAC	GTGGTCCTGG	GGGATTTTGA	AGATGTTAAA	1680
	GGGAAAAGAT	GACAGAAAGA	AAAGTATACG	AGAGAAACCT	AAAGTCTCTG	ACTCAGACAA	1740
	TAATGAAGGT	TCATCTTTCC	CTGCTCCTCC	TAAACAATTG	GACATGGGAG	ATGAAGTTTA	1800
30	CGATGATGTG	GATACCTCTG	ATTTCCCTGT	TTCATCAGCA	GAGATGAGTC	AAGGAACTAA	1860
		GCTAAGACAG	`		1 '		1920
35	ARAAAAAGAC	TTCAGGAAAA	AATTTAAATA	TGATGGTGAA	ATTAGAGTCC	TATATTCAAC	1980
		ACTTCCATAA					2040
4.0		CTAGAAGTTA					2100
40		TATGGTTATG					2160
		GCTGATGGCT					2220
45		TTAGGTGCCA					2280
		ATGCACAAAA					2340
50		AAAGTTTGAA					2400
50		TGAGATGTAA					2460
		AAGGTACTTC	•			AAAAAAAA	2520
55	AAAAAAAAA	AAAAAAAAA	ACTCGAGGGG	GGGCCCGGTA	CCCAAT		2566

⁽²⁾ INFORMATION FOR SEQ ID NO: 88:

	(1) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 540 base pairs	
	(B) TYPE: nucleic acid (C) STRANDEDNESS: double	
5	(D) TOPOLOGY: linear	•
,	(b) forobost. Thier	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 88:	
10	GAATTCGGCA CGAGGCTTTC TGTGTCCTCT GTGGCTGCTT TAGTGTGCCA CCAGGGGCAG	60
	ACTTGGGTGG GTTGCAGCAG AGATGGCATG GCCCTCAAGG TCCAAGATGT TTACTCTCTT	120
	GCCGGTCCTC TGTTATCTCT GGTCTTTGTG GTTGCCACAG TTTTCTTGGA TCCAGGAGTT	180
15	AAAGGCAGTC CTGAGGGATG ATGGCCTCAT CTCCGCAGTT GCYTGGAATG CTGAATTTCA	240
	GACGTGCTAA AGGAGGGTTG CAGACATTGT GTGGWATGCA TTCAGACCCC AGATGTGGGT	300
20	GCAGGAAGGC AGGCATGGCA CAGCCAGGTA GAGACTGGTT TCCAGGCCCA AGCAGCCTTC	360
	AGCAGCTGTG CGCCTTGTTT CTGATGTTGT TTGGGAGTAA GAATAATGTA GACATGGGG	420
	GTCATGARGC TCAATAAAAA CTTCAAGGAA ACCTCCCATG GCATGGTTGG GCGCAGTGAC	480
25	TCATGCCTGT AACCCCAGCA CTGTGGAATG CCAAGGTGGA AGGATCGCTT GAGGCCAAGA	540
30 35	(2) INFORMATION FOR SEQ ID NO: 89: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1863 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 89:	
40	TCGACCCACG CGTCCGGCGA GATCCCTACC GCAGTAGCCG CCTCTGCCGC CGCGGAGCTT	60
	CCCGAACCTC TTCAGCCGCC CGGAGCCGCT CCCGGAGCCC GGCCGTAGAG GCTGCAATCG	120
45	CAGCCGGGAG CCCGCAGCCC GCGCCCCGAG CCCGCCGCCG CCCTTCGAGG GCGCCCCAGG	180
	CCGCGCCATG GTGAAGGTGA CGTTCAACTC CGCTCTGGCC CAGAAGGAGG CCAAGAAGGA	240
	CGAGCCCAAG AGCGGCGAGG AGGCGCTCAT CATCCCCCCC GACGCCGTCG CGGTGGACTG	300
50	CAAGGACCCA GATGATGTGG TACCAGTTGG CCAAAGAAGA GCCTGGTGTT GGTGCATGTG	360
	CTTTGGACTA GCATTTATGC TTGCAGGIGT TATTCTAGGA GGAGCATACT TGTACAAATA	420
55	TTTTGCACTT CAACCAGATG ACGTGTACTA CTGTGGAATA AAGTACATCA AAGATGATGT	480
	CATCTTAAAT GAGCCCTCTG CAGATGCCCC AGCTGCTCTC TACCAGACAA TTGAAGAAAA	540
	TATTAAAATC TITGAAGAAG AAGAAGTIGA ATTTATCAGT GIGCCIGICC CAGAGITIGC	600

AGATAGTGAT CCTGCCAACA TTGTTCATGA CTTTAACAAG AAACTTACAG CCTATTTAGA

120

	TCTTAACCTG	GATAAGTGCT	ATGTGATCCC	TCTGAACACT	TCCATTGTTA	TGCCACCCAG	720
5	AAACCTACTG	GAGTTACTTA	TTAACATCAA	GCTGGAACC	TATTTGCCTC	AGTCCTATCT	780
,	GATTCATGAG	CACATGGTTA	TTACTGATCG	CATTGAAAAC	ATTGATÇACC	TGGGTTTCTT	840
	TATTTATCGA	CTGTGTCATG	ACAAGGAAAC	TTACAAACTG	CAACGCAGAG	AAACTATTAA	900
10	AGGTATTCAG	AAACGTGAAG	CCAGCAATTG	TTTCGCAATT	CGGCATTITG	AAAACAAATT	960
	TGCCGTGGAA	ACTITAATIT	GPTCTTGAAC	AGTCAAGAAA	AACATTATTG	AGGAAAATTA	1020
15	ATATCACAGC	ATAACCCCAC	CCTTTACATT	TTGTGCAGTG	TTTTTTATTA	AAAGTCTTCT	1080
1.5	TTCATGTAAG	TAGCAAACAG	GGCTTTACTA	TCTTTTCATC	TCATTAATTC	AATTAAAACC	1140
	ATTACCTTAA	AATTTTTTC	TTTCGAAGTG	TGGTGTCTTT	TATATTTGAA	TTAGTAACTG	1200
20	TATGAAGTCA	TAGATAATAG	TACATGTCAC	CTTAGGTAGT	AGGAAGAATT	ACAATTTCTT	1260
	TAAATCATTT	ATCTGGATTT	TTATGTTTTA	TTAGCATTTT	CAAGAAGACG	GATTATCTAG	1320
25	AGAATAATCA	TATATATGCA	TACGTAAAAA	TGGACCACAG	TGACTTATTT	GTAGTTGTTA	1380
	GTTGCCCTGC	TACCTAGTTT	GTTAGTGCAT	TTGAGCACAC	ATTTTAATTT	TCCTCTAATT	1440
	AAAATGTGCA	GTATTTTCAG	TGTCAAATAT	ATTTAACTAT	TTAGAGAATG	ATTTCCACCT	1500
30	TTATGTTTTA	ATATCCTAGG	CATCTGCTGT	AATAATATTT	TAGAAAATGT	TTGGAATTTA	1560
	AGAAATAACT	TGTGTTACTA	ATTTGTATAA	CCCATATCTG	TGCAATGGAA	TATAAATATC	1620
35	ACAAAGTTGT	TTAACTAGAC	TGCGTGTTGT	TTTTCCCGTA	TAATAAAACC	AAAGAATAGT [']	1680
_	TTGGTTCTTC	AAATCTTAAG	AGAATCCACA	TAAAAGAAGA	AACTATTTT	TAAAAATTCA	1740
	CTTCTATATA	. TACAATGAGT	AAAATCACAG	ATTITTTCTT	AAAAAAT	TAAGTCATTT	1800
40	ТААТААСТАА	ACCAGATTCT	TTGTGATACT	ATTAANGTAA	CATTTAGCCC	САААААААА	1860
	AAA						1863
45							
	(2) INFORM	NATION FOR S	EQ ID NO: 9	0:			
	(1)	SEQUENCE C	HARACTERISI	TCS:			
50	(1)	-	IGIH: 2478 l				
		• •	E: nucleic				
		• - •	RANDEDNESS: POLOGY: line				
		(D) 101					

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 90:

GCCACAGCGG CACGAGGTGA GCTGAGCCGG TGGGTGAGCG GCGGCCACGG CATCCTGTGC

TGTGGGGGT ACGAGGAAAG ATCTAATTAT CATGGACCTG CGACAGTTTC TTATGTGCCT

BNSDOCID: <WO___9842738A1_I_>

55

	GICCCIGIGO	ACAGCCTTTG	CCTTGAGCAA	ACCCACAGAA	AAGAAGGACC	GTGTACATCA	180
	TGAGCCTCAG	CTCAGTGACA	AGGTTCACAA	TGATGCTCAG	AGTTTTGATT	ATGACCATGA	240
5	TGCCTTCTTG	GGTGCTGAAG	AAGCAAAGAC	CTTTGATCAG	CTGACACCAG	AAGAGAGCAA	300
	GGAAAGGCTT	GGAAAGATTG	TAAGTAAAAT	AGATGGCGAC	AAGGACGGGT	TTGTCACTGT	360
10	GGATGAGCTC	AAAGACTGGA	TTAAATTTGC	ACAAAAGCGC	TGGATTTACG	AGGATGTAGA	420
10	GCGACAGTGG	AAGGGGCATG	ACCTCAATGA	GGACGGCCTC	GITTCÇTGGG	AGGAGTATAA	480
	AAATGCCACC	TACGGCTACG	TTTTAGATGA	TCCAGATCCT	GATGATGGAT	TTAACTATAA	540
15	ACAĢATGATG	GTTAGAGATG	ACCGGAGGTT	TAAAATGGCA	GACAAGGATG	GAGACCTCAT	600
	TGCCACCAAG	GAGGAGTTCA	CAGCTTTCCT	GCACCCTGAG	GAGTATGACT	ACATGAAAGA	660
20	TATAGTAGTA	CAGGAAACAA	TGGAAGATAT	AGATAAGAAT	GCTGATGGTT	TCATTGATCT	720
20	AGAAGAGTAT	ATTGGTGACA	TGTACAGCCA	TGATGGGAAT	ACTGATGAGC	CAGAATGGGT	780
	AAAGACAGAG	CGAGAGCAGT	TIGITGAGTT	TCGGGATAAG	AACCGTGATG	GGAAÇATGGA	840
25	CAAGGAAGAG	ACCAAAGACT	GGATCCTTCC	CTCAGACTAT	GATCATGCAG	AGGCAGAAGC	900
	CAGGCACCTG	GTCTATGAAT	CAGACCAAAA	CAAGGATGGC	AAGCTTACCA	AGGAGGAGAT	960
30	CGTTGACAAG	TATGACTTAT	TTGTTGGCAG	CCAGGCCACA	GATTTTGGGG	AGGCCTTAGT	1020
	ACGGCATGAT	GAGTTCTGAG	CTRCGGAGGA	ACCCTCATTT	CCTCAAAAGT	AATTTATTTT	1080
	TACAGCTTCT	GGTTTCACAT	GAAATTGTTT	GCGCTACTGA	GACTGTTACT	ACAAACTTTT	1140
35	TAAGACATGA	AAAGGCGTAA	TGAAAACCAT	CCCGTCCCCA	TTCCTCCTCC	TCTCTGAGGG	1200
	ACTGGAGGGA	AGCCGTGCTT	CTGAGGAACA	ACTCTAATTA	GTACACTTGT	GTTTGTAGAT	1260
40	TTACACTTTG	TATTATGTAT	TAACATGGCG	TGTTTATTTT	TGTATTTTC	TCTGGTTGGG	1320
	AGTATGATAT	GAAGGATCAA	GATCCTCAAC	TCACACATGT	AGACAAACAT	TAGCTCTTTA	1380
	CTCTTTCTCA	ACCCCTTTTA	TGATTTTAAT	AATTCTCACT	TAACTAATTT	TGTAAGCCTG	1440
45	AGATCAATAA	GAAATGTTCA	GGAGAGAGGA	AAGAAAAAA	ATATATGCTC	CACAATTTAT	1500
	ATTTAGAGAG	AGAACACTTA	GTCTTGCCTG	TCAAAAAGTC	CAACATTTCA	TAGGTAGTAG	1560
50	GGCCACATA	TTACATTCAG	TTGCTATAGG	TCCAGCAACT	GAACCTGCCA	TTACCTGGGC	1620
	AAGGAAAGAT	CCCTTTGCTC	TAGGAAAGCT	TGGCCCAAAT	TGATTTTCTT	CTTTTTCCCC	1680
	CTGTAGGACT	GACTGTTGGC	TAATTTTGTC	AAGCACAGCT	GTGGTGGGAA	GAGTTAGGGC	1740
55	CAGTGTCTTG	AAAATCAATC	AAGTAGTGAA	TGTGATCTCT	TTGCAGAGCT	ATAGATAGAA	1800
	ACAGCTGGAA	AACTAAAGGA	AAAATACAAG	TGTTTTCGGG	GCATACATTT	TTTTTCTGGG	1860
50	TGTGCATCTG	TIGAAATGCT	CAAGACTTAA	TTATTTGCCT	TTTGAAATCA	CTGTAAATGC	1920

	CCCCATCCGG	TTCCTCTTCT	TCCCAGGTGT	GCCAAGGAAT	TAATCTTGGT	TTCACTACAA	1980
	TTAAAATTCA	CTCCTTTCCA	ATCATGTCAT	TGAAAGTGCC	TTTAACGAAA	GAAATGGTCA	2040
5	CTGAATGGGA	ATTCTCTTAA	GAAACCCTGA	GATTAAAAAA	AGACTATITG	GATAACTTAT	2100
	AGGAAAGCCT	AGAACCTCCC	AGTAGAGTGG	GGATTTTTTT	CTTCTTCCCT	TTCTCTTTTG	2160
0	GACAATAGTT	AAATTAGCAG	TATTAGTTAT	GAGTTTGGTT	GCAGTGTTCT	TATCTTGTGG	2220
.0	GCTGATTTCC	AAAAACCACA	TGCTGCTGAA	TTTACCAGGG	ATCCTCATAC	CTCACAATGC	2280
	AAACCACTTA	CTACCAGGCC	TTTTTCTGTG	TCCACTGGAG	AGCTTGAGCT	CACACTCAAA	2340
5	GATCAGAGGA:	CCTACAGAGA	GGGCTCTTTG	GTTTGAGGAC	CATGGCTTAC	CITTCCIGCC	2400
	TTTGACCCAT	CACACCCCAT	TICCICCICT	TTCCCTCTCC	CCGCTGCCAA	TTCCTGCAGC	2460
20	CCGGGGGAAC	CACTAGTT		•			2478

(2) INFORMATION FOR SEQ ID NO: 91:

25

30

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2058 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 91:

35	TCGGCCTTGC	TTTTGTGGYC	TICCICIGIG	GCCAGAGCGT	TTTCATCACC	AAGCCTCCTG	60
55	ATGGCAGTNC	CTTCACCGAT	ATGTTCAAGA	TACTGACGTA	TTCCTGCTGT	TCCCAGAAGC	120
	GAAGTGGAGA	GCGCCAGAGT	AATGGTGAAG	GCATTGGAGT	NTTTCAGCAA	TCTTCTAAAC	180
40	AAAGTCTGTT	TGATTCATGT	AAGATGTCTC	ATGGTGGGCC	ATTTACAGAA	GAGAAAGTGG	240
	AAGATGTGAA	AGCTCTGGTC	AAGATTGTCC	CIGITITCIT	GCTTTGATA	CCTTACTGGA	300
45	CAGTGTATTT	CCAAATGCAG	ACAACATATG	TTTTACAGAG	TCTTCATTTG	AGGATTCCAG	360
43	AAATTTCAAA	TATTACAACC	ACTCCTCACA	CGCTCCCTGC	AGCCTGGCTG	ACCATGTTTG	420
	ATGCTGTGCT	CATCCTCCTG	CTCATCCCTC	TGAAGGACAA	ACTGGTCGAT	CCCATTTTGA	480
50	GAAGACATGG	CCTGCTCCCA	TCCTCCCTGA	AGAGGATCGC	CGTGGGCATG	TTCTTTGTCA	540
	TGTGCTCRGC	CTTTGCTGCA	GGAATTTTGG	AGAGTAAAAG	GCTGAACCTT	GTTAAAGAGA	600
55	AAACCATTAA	TCAGACCATC	GGCAACGTCG	TCTACCATGC	TGCCGATCTG	TCCCTCTCCT	660
33	GGCAGGTGCC	GCAGTACTTG	CTGATTGGGA	TCAGCGAGAT	CTTTGCAAGT	ATCGCAGGCC	720
	TGGAATTTGC	ATACTCAGCT	GCCCCCAAGT	CCATGCAGAG	TGCCATAATG	GCCTTCTTCT	780
60	TTTTCTTCTC	TGGCGTCGGG	TCGTTCGTGG	GTTCTGGACT	GCTGGCACTG	GIGICTATCA	840

	AAGCCATCGG	ATGGATGAGC	AGTCACACAG	ACTITGGTAA	TATTAACGGC	TGCTATTTGA	900
5	ACTATTACTT	TTTCCTTCTG	GCTGCTATTC	AAGGAGCTAC	CCTCCTGCTT	TŢĊĊŢĊĀŢŢĀ	960
	TTTCTGTGAA	ATATGACCAT	CATCGAGACC	ATCAGCGATC	AAGAGCCAAT	GCCGTGCCCA	1020
	CCAGCAGGAG	GGCCTGACCT	TCCTGAGGCC	ATGTGCGGTT	TCTGAGGCTG	ACATGTCAGT	1080
10	AACTGACTGG	GGTGCACTGA	GAACAGGCAA	GACTTTAAAT	TCCCATAAAA	TGTCTGACTT	1140
	CACTGAAACT	TGCATGTTGC	CTGGATTGAT	TTCTTCTTTC	CCTCTATCCA	AAGGAGCTTG	1200
15	GTAAGTGCCT	TACTGCAGCG	TGTCTCCTGG	CACGCTGGGC	CCTCCGGGAG	GAGAGCTGCA	1260
10	GATTTCGAGT	ATGTCGCTTG	TCATTCAAGG	TCTCTGTGAA	TCCTCTAGCT	GGGTTCCCTT	1320
	TTTTACAGAA	ACTCACAAAT	GGAGATTGCA	AAGTCTTGGG	GAACTCCACG	TGTTAGTTGG	1380
20	CATCCCAGTT	TCTTAAACAA	ATAGTATCAC	CTGCTTCCCA	TAGCCATATC	TCACTGTAAA	1440
	TTAAAAAAA	AATAAACTGT	TACTTATATT	TAAGAAAGTG	AGGATTTTTT	TTTTTTAAAG	1500
25	ATAAAAGCAT	GGTCAGATGC	TGCAAGGATT	TTACATAAAT	GCCATATTTA	TGGTTTCCTT	1560
	CCTGAGAACA	ATCTTGCTCT	TGCCATGPTC	TTTGATTTAG	GCTGGTAGTA	AACACATTTC	1620
	ATCTGCTGCT	TCAAAAAGTA	CTTACTTTTT	AAACCATCAA	CATTACTTTT	CTTTCTTAAG	1680
30	GCAAGGCATG	CATAAGAGTC	ATTTGAGACC	ATGTGTCCCA	TCTCAAGCCA	CAGAGCAACT	1740
	CACGGGGTAC	TTCACACCTT	ACCTAGTCAG	AGTGCTTATA	TATAGCTTTA	TTTTGGTACG	1800
35	ATTGAGACTA	AAGACTGATC	ATGGTTGTAT	GTAAGGAAAA	CATTCTTTTG	AACAGAAATA,	1860
	GTGTAATTAA	AAATAATTGA	AAGTGTTAAA	TGTGAACTTG	AGCTGTTTGA	CCAGTCACAT	1920
	TTTTGTATTG	TTACTGTACG	TGTATCTGGG	CCTTCTCCCT	TTGTTAATAC	TTTTTCTGTA	1980
40	TTTGTTGCTG	TATTTTTGGC	ATAACTTTAT	TATAAAAAGC	ATCTCAAATG	CGAAAWAAAA	2040
	АААААААА	AAAAAAAC					2058
45							
	(2) INFORM	ATION FOR SE	EO ID NO: 92	1 .			
		SEQUENCE CI	_				
50	(-,	(A) LEN	3TH: 1411 b	ase pairs			
		(C) STR	E: nucleic a ANDEDNESS: a	double			
		(D) TOP	OLOGY: line	ar			
55	ix)) SEQUENCE I	DESCRIPTION	: SEQ ID NO	: 92:		
	GGCACAGGAG	CGACCCGGGA	GAAGGAGGC	CAMGAKGCGG	AAGCGGAGGA	GTCTCCAGGA	60
	GACCCGGGGA	CAGCATCGCC	CAGGCCCCTG	TTTGCAGGCC	TTTCAGATAT	ATCCATCTCA	120

	CAAGACATCC CCGTAGAAGG AGAAATCACC ATTCCTATGA GATCTCGCAT CCGGGAGTTT	180
	GACAGCTCCA CATTAAATGA ATCTGTTCGC AATACCATCA TGCGTGATCT AAAAGCTGTT	240
5	GGGAAAAAT TCATGCATGT TITGTACCCA AGGAAAAGTA ATACTCTTTT GAGAGATTGG	300
	GATTTGTGGG GCCCTTTGAT CCTTTGTGTG ACACTCGCAT TAATGCTGCA AAGAGACTCT	360
10	GCAGATAGTG AAAAAGATGG AGGCCCCCAA TTTGCAGAGG TGTTTGTCAT TGTCTGGTTT	420
	GGTGCAGTTA CCATCACCCT CAACTCAAAA CTTCTTGGAG GGAACATATC TTTTTTTCAG	480
	AGCCTCTGTG TGCTGGGTTA CTGTATACTT CCCTTGACAG TAGCAATGCT GATTTGCCGG	540
15	CTGGTACTTT TGGCTGATCC AGGACCTGTA AACTTCATGG TTCGGCTTTT TGTGGTGATT	600
	GTGATGTTTG CCTGGTCTAT AGTTGCCTCC ACAGCTTTCC TTGCTGATAG CCAGCCTCCA	660
20	AACCGCAGAG CCCTAGCTGT TTATCCTGTT TTCCTGTTTT ACTTTGTCAT CAGTTGGATG	720
	ATTCTCACCT TTACTCCTCA GTAAATCAGG AATGGGAAAT TAAAAACCAG TGAATTGAAA	780
	GCACATCTGA AAGATGCAAT TCACCATGGA GCTTTGTCTC TGGCCCTTAT TTGTCTAATT	840
25	TTGGAGGTAT TTGATAACTG AGTAGGTGAG GAGATTAAAA GGGAGCCATA TAGCACTGTC	900,
	ACCCCTTATT TGAGGAACTG ATGTTTGAAA GGCTGTTCTT TTCTCTCTTA ATGTCATTTC	960
30	TTTAAAAATA CATGTGCATA CTACACAGG TATATAATGC CTCCTTAAGG CATGATGGAG	1020
	TCACCGTGGT CCATTTGGGT GACAACCAGT GACTTGGGAA GCACATAGAT ACATCTTACA	1080
	AGITGAATAG AGITGATAAC TATTITCAGI TITGAGAATA CCAGITCAGG TGCAGCICIT	1140
35	AAACACATTG CCTTATGACT ATTAGAATAT GCCTCTCTTT TCATAAATAA AAATACATGG	1200
	TCTATATCCA TTTTCTTTTA TTTCTCTCTC TTAAGCTTAA AAAGGCAATG AGAGAGGTTA	1260
40	GGAGTGGGTT CATACACGGA GAATGAGAAA ACATGCATTA ACCAATATTC AGATTTTGAT	1320
	CAGGGGAAAT TCTAYACTTG TTGCAAAAAA AAAAAAAAAA AAACTCGAGG GGGGCCCGGT	1380
	ACCCAATCGC NGTATATGAT CGNAAACAAT C	1411
45		
	(2) INFORMATION FOR SEQ ID NO: 93:	
50	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 2187 base pairs (B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double (D) TOPOLOGY: linear	
55	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 93:	
	GCTTTGGCTT TTTTTGGCGG ACTGGGGCGC CCTCCGGAAG CGTTTCCAAC TTTCCAGAAG	60
60	TTTCTCGGGA CGGGCAGGAG GGGGTGGGGA CTGCCATATA TAGATCCCGG GAGCAGGGGA	60 120
	The second secon	120

	GCGGGCTAAG	AGTAGAATCG	TGTCGCGCTC	GAGAGCGAGA	GTCACGTCCC	GCCCTACCC	180
5	CAGCCCGACC	CAGGCCCACC	GTGGTGCACG	CAAACCACTT	.CCTGGCCATG	CCCTCCCTCC	240
,	TGCTTCTCAG	CGCCTTCTGC	CTCCTGGAGG	CGGCCCTGGC	CGCCGAGGTG	AAGAAACCTG	300
	CAGCCGCAGC	AGCTCCTGGC	ACTGCGGAGA	AGTTGAGCCC	CAAGGCGGCC	ACGCTTGCCG	360
10	AGCGCAGCCG	GCCIGGCCIT	CAGCTTGTAC	CAGGCCATGG	CCAAGGACCA	GCAGTGGAG	420
	AACATCCTGG	TGTCACCCGT	GGTGGTGGCC	TCGTCGCTGG	GCTCGTGTC	GCTGGGCGGC	480
15	AAGGCGACCA	CGGCGTCGCA	GGCCAAGGCA	GTGCTGAGCG	CCGAGCAGCT	GCGCGACGAG	540
	GAGGTGCACG	CCGCCTGGG	CGACCTCCTG	CGCTCACTCA	GCAACTCCAC	GGCGCGCAAC	600
	GTGACCTGGA	AGCTGGGCAG	CCGACTGTAC	GGACCCAGCT	CAGTGAGCTT	CGCTGATGAC	660
20	TTCGTGCGCA	GCAGCAAGCA	GCACTACAAC	TGCGAGCACT	CCAAGATCAA	CTTCCGCGAC	720
	AAGCGCAGCG	CGCTGCAGTC	CATCAACGAG	TGGGCCGCGC	AGACCACCGA	CGGCAAGCTG	780
25	CCCGAGGTCA	CCAAGGACGT	GGAGCGCACG	GACGGCGCCC	TGTTAGTCAA	CCCCATCTTC	840
	TTCAAGCCAC	ACTGGGATGA	GAAATTCCAC	CACAAGATGG	TGGACAACCG	TGGCTTCATG	9,00
	GTGACTCGGT	CCTATACCGT	GGGTGTCATG	ATGATGCACC	GGACAGGCCT	CTACAACTAC	960
30	TACGACGACG	AGAAGGAAAA	GCTGCAAATC	GTGGAGATGC	CCCTGGCCCA	CAAGCTCTCC	1020
	AGCCTCATCA	TCCTCATGCC	CCATCACGTG	GAGCCTCTCG	AGCGCCTTGA	AAAGCTGCTA	1080
35	ACCAAAGAGC	AGCTGAAGAT	CTGGATGGGG	AAGATGCAGA	AGAAGGCTGT	TGCCATCTCC	1140
	TTGCCCAAGG	GTGTGGTGGA	GGTGACCCAT	GACCTGCAGA	AACACCTGGC	TGGGCTGGGC	1200
	CTGACTGAGG	CCATTGACAA	GAACAAGGCC	GACTTGTCAC	GCATGTCAGG	CAAGAAGGAC	1260
40	CTGTACCTGG	CCAGCGTGTT	CCACGCCACC	GCCTTTGAGT	TGGACACAGA	TGGCAACCCT	1320
	TTGACCAGAA	TTACGGGCGG	AGGAGTGCGC	ACCCAAGTGT	TCTACGCCGA	CCACCCCTTC	1380
45	ATTTCCTAGT	GCGGGACACC	CAAAGCGGTC	CCTGCTATTC	ATTGGGCGCC	TECTCCECC	1440
	TAAGGGTGAC	AAGATGCGAG	ACGAGTTATA	GCCTCAGGG	TGCACACAGG	ATGGCAGGAG	1500
	GCATCCAAAG	GCTCCTGAGA	CACATGGGTG	CTATTGGGGT	TGGGGGGAG	GTGAGGTACC	1560
50	AGCCTTGGAT	ACTCCATGGG	GTGGGGTGGA	AAAGCAGACC	GGGTTCCCG	TETECCTGAG	1620
	CGGACTTCCC	AGCTAGAATT	CACTCCACTT	GGACATGGGC	CCCAGATACC	ATGATGCTGA	1680
55	GCCCGGAAAC	TCCACATCCT	GTGGGACCTG	GGCCATAGTC	ATTCTGCCTG	CCCTGAAAGT	1740
	CCCAGATCAA	GCCTGCCTCA	ATCAGTATTC	ATATTTATAG	CCAGGTACCT	TCTCACCTGT	1800
	GAGACCAAAT	'TGAGCTAGGG	GGGTCAGCCA	GCCCTCTTCT	GACACTAAAA	CACCTCAGCT	1860
60	GCCTCCCCAG	CTCTATCCCA	ACCTCTCCCA	ACTATAAAAC	TAGGTGCTGC	AGCCCCTGGG	1920

	ACCAGGCACC CCCAGAATGA CCTGGCCGCA GTGAGGCGGA TTGAGAAGGA GCTCCCAGGA	1980
5	GGGGCTTCTG GGCAGACTCT GGTCAAGAAG CATCGTGTCT GGCGTTGTGG GGATGAACTT	2040
_	TTTGTTTTGT TTCTTCCTTT TTTAGTTCTT CAAAGATAGG GAGGGAAGGG GGAACATGAG	2100
	CCTTGTTGC TATCAATCCA AGAACTTATT TGTACATTTT TTTTTTCAAT AAAACTTTTC	2160
10	CAATGACAAA AAAAAAAA AAAAAAA	2187
15	(2) INFORMATION FOR SEQ ID NO: 94:	ı
20	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 757 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 94:	
25	GACAGTACGG TCGGATTCCC GGGTCGACCC ACGCGTCCGC GGACGGTGAA GAAGGTGAAG	60
	ATGCCGTGG CCAGGGCCGG GGTCTTGGGA GTCCAGTGGC TGCAAAGGGC ATCCCGGAAC	120
30	GTGATGCCGC TGGGCGCACG GACAGCCTCC CACATGACCA AGGACATGTT CCCGGGGCCC	180
	TATCCTAGGA CCCCAGAAGA ACGGGCCGCC GCCGCCAAGA AGTATAATAT GCGTGTGGAA	240
	GACTACGAAC CTTACCCGGA TGATGGCATG GGGTATGGCG ACTACCCGAA GCTCCCTGAC	300
35	CGCTCACAGC ATGAGAGAGA TCCATGGTAT AGCTGGGACC AGCCGGGCCT GAGGTTGAAC	360
	TGGGGTGAAC CGATGCACTG GCACCTAGAC ATGTACAACA GGAACCGTGT GGATACATCC	420
40	CCCACACCTG TTTCTTGGCA TGTCATGTGT ATGCAGCTCT TCGGTTTCCT GGCTTTCATG	480
	ATATTCATGT GCTGGGTGGG GGACGTGTAC CCTGTCTACC AGCCTGTGGG ACCAAAGCAG	540
	TATCCTTACA ATAATCTGTA CCTGGAACGA GGCGGTGATC CCTCCAAAGA ACCAGAGCGG	600
1 5	GIGGITCACT ATGAGATCIG AGGAGGCTIC GIGGGCTITT GGGTCCTCTA ACTAGGACTC	660
	CCTCATTCCT AGAAATTTAA CCTTAATGAA ATCCCTAATA AAACTCAGTG CTGTGTTAAA	720
50	AAAAAAAAA AAAAAAAAA AAAAAGGGGG GCCCCNN	757
55	(2) INFORMATION FOR SEQ ID NO: 95:	
	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 2394 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double	
50	(D) TOPOLOGY: linear	

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 95:

.5	GGCACGAGCA CTCCTGCACT TCCCCACCCC CACGACCGAA CCTGGCTTCG CTAACGCCCT	60
3	CCCAGCTCCC TCGGGCCTGA CTTCCGGTTT CCTCGCGCGT CCCTGGCGCC GAGCCGCGGA	120
	CAGCAGCCCC TTTTCCGGCT GAGAGCTCAT CCACACTTCC AATCACTTTC CGGAGTGCTT	180
10	cocciocate deecccese sessectes estateses ecaintiget	240
	GOGTAACGGG CCTTCTCYCG CGTCGGCCCG GCCCCTCCTG CCTCGGCCTCG TCCCTCCTTC	300
15	CAGAACGTCC CGGGCTCCTG CCGAGTCAGA AGAAATGGGA CTCCCTCCGC GACGTGCCCG	360
13	GAGCAGCTCC CTTCGCTGTG GAAGCGGCGG TGTCTTCGAA GAAACCGGAA GCCCGTGGTG	420
	ACCCCTGGCG ACCCGGTTTG TTTTCGGTCC GTTTCCAAAC ACTAAGGAAT CGAAACTCGG	480
20	COSCCTTGGG GGCGGCCCTA CGTAGCCTGG CTTCTGGTTG TCATGGATGC ACTGGTAGAA	540
	GATGATATCT GTATTCTGAA TCATGAAAAA GCCCATAAGA GAGATACAGT GACTCCAGTT	600
25	TCAATATATT CAGGAGATGA ATCTGTTGCT TCCCATTTTG CTCTTGTCAC TGCATATGAA	660
23	GACATCAAAA AACGACTTAA GGATTCAGAG AAAGAGAACT CTTTGTTAAA GAAGAGAATA	720
	AGATTITIGG AAGAAAAGCT AATAGCTCGA TITGAAGAAG AAACAAGTTC CGTGGGACGA	780
30	GAACAAGTAA ATAAGGCCTA TCATGCATAT CGAGAGGTTT GCATTGATAG AGATAATTTG	840
	AAGAGCAAAC TGGACAAAAT GAATAAAGAC AACTCTGAAT CITTGAAAGT ATTGAATGAG	900
35	CAGCTACAAT CTAAAGAAGT AGAACTCCTC CAGCTGAGGA CAGAGGTGGA AACTCAGCAG	960
55	GTGATGAGGA ATTTAAATCC ACCTTCATCA AACTGGGAGG TGGAAAAGTT GAGCTGTGAC	1020
	CTGAAGATCC ATGGTTTGGA ACAAGAGCTG GAACTGATGA GGAAAGAATG TAGCGATCTC	1080
40	AAAATAGAAC TACAGAAAGC CAAACAAACG GATCCATATC AGGAAGACAA TCTGAAGAGC	1140
	AGAGATCTCC AAAAACTAAG CATTTCAAGT GATAATATGC AGCATGCATA CTGGGAACTG	1200
45	AAGAGAGAAA TGTCTAATTT ACATCTGGTG ACTCAAGTAC AAGCTGAACT ACTAAGAAAA	1260
15	CTGAAAACCT CAACTGCAAT CAAGAAAGCC TGTGCCCCTG TAGGATGCAG TGAAGACCTT	1320
	GGAAGAGACA GCACAAAACT GCACTTGATG AATTTTACTG CAACATACAC AAGACATCCC	1380
50	CCTCTCTTAC CAAATGGCAA AGCTCTTTGT CATACCACAT CTTCCCCTTT ACCAGGAGAT	1440
	GTAAAGGTTT TATCAGAGAA AGCAATCCTC CAATCATGGA CAGACAATGA GAGATCCATT	1500
55	CCTAATGATG GTACATGCTT TCAGGAACAC AGTTCTTATG GCAGAAATTC TCTGGAAGAC	1560
	AATTCCTGGG TATTTCCAAG TCCTCCTAAA TCAAGTGAGA CAGCATTTGG GGAAACTAAA	1620
	ACTAAAACIT TGCCTTTACC CAACCTTCCA CCACTGCATT ACTTGGATCA ACATAATCAG	1680
60	AACTGCCTTT ATAAGAATTA ATTTGGAAGA GATTCACGAT TTCACCATGA GGACACTTAT	1740

	CTCTTTCAGT GGTCCTCCCA AGAAATTATT TAACAAACTG AANGGAGATT TTGATTAAAA	1800
5	TTTTGCAGAG GTCTTCAGTA TCTATATTTG AACACACTGT ACAATAGTAC AAAAACCAAC	1860
J	ATAGTTGGTT TTCTAGTATG AAAGAGCACC CTCTAGCTCC ATATTCTAAG AATCTGAAAT	1920
	ATGCTACTAT ACTAATTAAT AAGTAAACTT AAGGTGTTTA AAAAACTCTG CCTTCTATAT	. 1980
10	TAATTGTAAA ATTTTGCCTC TCAGAAGAAT GGAATTGGAG ATTGTAGACG TGGTTTTACA	2040
	AAATGIGAAA TGICTAAATA TCIGITCATA AAAATAAAAG GAAAACATGI TTCITCAAAT	2100
15	TGCATAATGG AACAAATGGC AATGTGAGTA GGTTACATTT CTGTTGTTAT AATGCGTAAA	2160
	GATATTGAAA ATATAATGAA ATAAAAGCAT CTTAGGTTAT ACCATCTTTA TATGCTATTG	2220
	CGITTCAATA TITAAGATTT AAAGIGATTT TITGGICACA GIGITTIGIT GATAAAATTT	2280
20	TTTTAGAATT GAAGITTGAA TTCTAAGACT TGAAACAACC TGATCACTGA AGCCAACTTT	2340
	GTCCCAGCAC ATTCCTTAAG TCCTAATTGG GGAAAAAAAA AAAAAAAAAC TCGA	2394
25 .		
	(2) INFORMATION FOR SEQ ID NO: 96:	
	(i) SEQUENCE CHARACTERISTICS:	
30	(A) LENGTH: 672 base pairs	
	(B) TYPE: nucleic acid (C) STRANDEDNESS: double	
	(D) TOPOLOGY: linear	
35		
-	(xi) SEQUENCE DESCRIPTION: SEQ ID NC: 96:	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NC: 96: AGTGCTCTGT TGCCCAGGCT GGAGTGCGTT AGTGTAATGT CAGTCCACTG CAACCTCCAC	60
40		60 120
	AGTGCTCTGT TGCCCAGGCT GGAGTGCGTT AGTGTAATGT CAGTCCACTG CAACCTCCAC	
	AGTGCTCTGT TGCCCAGGCT GGAGTGCGTT AGTGTAATGT CAGTCCACTG CAACCTCCAC CCCCAGGTTC AAGCAATTCT CATGCCTCAG CCTCCCAAGT AGCTGAAATT ACTGGCATGC	120
	AGTGCTCTGT TGCCCAGGCT GGAGTGCGTT AGTGTAATGT CAGTCCACTG CAACCTCCAC CCCCAGGTTC AAGCAATTCT CATGCCTCAG CCTCCCAAGT AGCTGAAATT ACTGGCATGC ACCACCACAC CCAGCTGATG TTTATTTATT TATTTATATA TTTATTTATT TTAGGTGTTT	120 180 240
40	AGTGCTCTGT TGCCCAGGCT GGAGTGCGTT AGTGTAATGT CAGTCCACTG CAACCTCCAC CCCCAGGTTC AAGCAATTCT CATGCCTCAG CCTCCCAAGT AGCTGAAATT ACTGGCATGC ACCACCACAC CCAGCTGATG TTTATTTATT TATTTTATATA TTTATTTTAT	120 180 240
40	AGTGCTCTGT TGCCCAGGCT GGAGTGCGTT AGTGTAATGT CAGTCCACTG CAACCTCCAC CCCCAGGTTC AAGCAATTCT CATGCCTCAG CCTCCCAAGT AGCTGAAATT ACTGGCATGC ACCACCACAC CCAGCTGATG TTTATTTATT TATTTATATA TTTATTTATT TTAGGTGTTT TTTTTTTTTT	120 180 240 300
40 45	AGTGCTCTGT TGCCCAGGCT GGAGTGCGTT AGTGTAATGT CAGTCCACTG CAACCTCCAC CCCCAGGTTC AAGCAATTCT CATGCCTCAG CCTCCCAAGT AGCTGAAATT ACTGGCATGC ACCACCACAC CCAGCTGATG TTTATTTATT TATTTATATA TTTATTTATT TTAGGTGTTT TTTTTTTTTT	120 180 240 300 360
40 45	AGTGCTCTGT TGCCCAGGCT GGAGTGCGTT AGTGTAATGT CAGTCCACTG CAACCTCCAC CCCCAGGTTC AAGCAATTCT CATGCCTCAG CCTCCCAAGT AGCTGAAATT ACTGGCATGC ACCACCACAC CCAGCTGATG TTTATTTATT TATTTATATA TTTATTTATT TTAGGTGTTT TTTTTTTTTT	120 180 240 300 360 420
40 45	AGTGCTCTGT TGCCCAGGCT GGAGTGCGTT AGTGTAATGT CAGTCCACTG CAACCTCCAC CCCCAGGTTC AAGCAATTCT CATGCCTCAG CCTCCCAAGT AGCTGAAATT ACTGGCATGC ACCACCACAC CCAGCTGATG TTTATTTATT TATTTATATA TTTATTTATT TTAGGTGTTT TTTTTTTTTT	120 180 240 300 360 420 480

60

CGTTAATTAC CC

(2) INFORMATION FOR SEQ ID NO: 97:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1419 base pairs

(B) TYPE: nucleic acid '

(C) STRANDEDNESS: double

10 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 97:

15	TAAGAACAGA	ACAGCAAGTA	TGAACCACAT	GGAACTTAAA	ACATATOGGT	GTGAAGTCCA	60
	CTTATGTAGA	CAAAACTTAT	AATTTCCAAA	CIGITGICIA	GTATACAGTG	ATCAGTTGCT	120
20	CTCTCTTCAA	GTCATTCCAC	ACATTTCCCT	ATTTTAGGCT	АТТАТААТАТ	AGAAAGAAAA	180
	TGGGAAGCAT	TAGTTGGAGC	TAGAAAATGA	ACTGTATATT	ATTGCTATAT	TTGCTAATAC	240
	CAACTATTTC	aataagtgtt	GTACCATATG	TAGCATTAAA	TATAAAATAC	ATAAAAGAAT	300
25	GTACAGAAAA	TAGCTTTTAT	TGAGTAATAT	TACATTTCAT	TTATACTGTA	GCAATATATT	360
	TGTAGGTATA	CTCTGTAAGG	GCTTTAAATA	AAAGAGGTCC	ATTAATACTT	CCTTATAAAA	420
	ATTCTAGTCT	GTTTCATTAC	TGCCCAGATG	TTTTAGAGAT	AAATATTTAT	GCAGAAGGTA	480
30	TTTTKGAAAG	TCYCCYTTTG	TCTGATAGAG	TTTAACNAGA	TATTTAAATT	TAGTGCYCNA	540
	GAAATCCCAC	AAGTCACGGT	CTAAACACAC	TTAGAATACT	ACAGCATAAA	TCTGTTAGCA	600
35	TTANTTGCCA	AATAAGACAG	TTGGGATCCC	AAACCCCAAG	TCCTTGAGCA	ATGTTTTTCC ,	660
	TCAAAAAGCT	GCTATNCCAA	TGATATAGGA	AAAWACATTG	TGTTTTCCTA	AACACACTTT	720
	TCTTTTTAAA	TGTGCTTCAT	TGTTTGATTT	GGTCCTGCCT	AAATTTCACA	AGCTAGGCCA	780
40	ATGAAGGCTG	AATCAAAGAC	ATTTCATCCA	CCAATATCAT	GTGTAGATAT	TATGTATAGA	840
	AAATAAAATA	AATTATGGCT	CTAACTTCTG	TGTTGCTGTT	TATCITGITA	TTTTTCGGCG	900
45	TTATACTAAT	GNGTTTATTG	AGAGCATTTT	ACCTTCCAGA	CTTCTCATGG	CTAACTTTTG	960
	GTCTGWATTT	TGSTCCTTAG	ATGKGAATAT	TTCTTATTAG	TYTGCTYCCT	GCWACGCAAT	1020
	GACTGCATTT	CTATCATTTC	TCAGTTTGTT	AGWATATGTG	GATAGTATTC	TACTGTATAA	1080
50	ATGATTGCAA	AGTTTATCAA	AAACAAATTA	TTATATGTAG	CTTTTCTACA	GIGCTITICCT	1140
	AAACCATGTA	GTACTAGTTA	AGTSTTCCTT	GAAAATAAAG	ATACACTCTT	ATAGGGGACA	1200
55	GITCCIGITC	ACTCCCAGGA	AACTTTTTTA	AAAGATGACA	CIGAATGITT	ATTGCACTTT	1260
	AGTGCAGTGA	AGTGGCAATA	AAACCTAACA	TGAATCAAGG	TTGTTTATGG	CAGATGCATG	1320
	TGTTGCTTTA	CAGAGTTTAG	CAAAAGCTCT	TAATTTTATG	TCATACTGTA	TTCTACTGAA	1380
60	TAATAAAGCT	AACATTATTC	AATAATAAA	TGGAAAAAA			1419

5 .	(2)	INFORMATION	FOR	SEO	ID	NO:	98:
,	14,	TALOMANITON	LOW	يوسر		110.	,,,

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1487 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 98:

					1		
15	GCGACCGCGC	CCCTTTCAGC	TAGCTCGCTC	GCTCGCTCTG	CTTCCCTGCT	GCCGGCTGCG	60
	CATGGCKWTG	CCCTTCCCCC	CCCTCCCCC	GGTCGAGCCG	GCCTGCGCAG	CCGGTACCAG	120
20	CAGTTGCAGA	ATGAAGAAGA	GTCTGGAGAA	CCTGAACAGG	CTCCAGGTGA	TGCTCCTCCA	180
	CCTTACAGCA	GCATTTCTGC	AGAGAGCGCA	GTTTTCCACC	TATTTCCCTG	GATATTTTGA	240
	TGGTCAGTAC	TGGCTCTGGT	GGCTGTTCCT	TGTTTTAGGC	TITCTCCTGT	TTCTCAGAGG	300
25	ATTTATCAAT	TATGCAAAAG	TTCGGAAGAT	GCCAGAAACT	TTCTCAAATC	TCCCCAGGAC	360
	CAGAGITCTC	TTTATTTATT	AAAGATGTTT	TCTGGCAAAG	GCCTTCCTGC	ATTTATGAAT	420
30	TCTCTCTCAA	GAAGCAAGAG	AACACCTGCA	GGAAGTGAAT	CAAGATGCAG	AACACAGAGG	480
	AATAATCACC	TGCTTTAAAA	AAATAAAGTA	CTGTTGAAAA	GATCATTTCT	CTCTATTIGT	540
	TCCTAGGTGT	AAAATTTTAA	TAGTTAATGC	AGAATTCTGT	AATCATTGAA	TCATTAGTGG	600
35	TTAATGTTTG	AAAAAGCTCT	TGCAATCAAG	TCTGTGATGT	ATTAATAATG	CCTTATATAT	660
	TGTTTGTAGT	CATTTTAAGT	AGCATGAGCC	ATGTCCCTGT	AGTCGGTAGG	GGGCAGTCTT	720
40	GCTTTATTCA	TCCTCCATCT	CAAAATGAAC	TTGGAATTAA	ATATTGTAAG	ATATGTATAA	780
40	TGCTGGCCAT	TTTAAAGGGG	TTTTCTCAAA	AGTTAAACTT	TTGTTATGAC	TGTGTTTTTG	840
	CACATAATCC	ATATTTCCTG	TTCAAGTTAA	TCTAGAAATT	TATTCAATTC	TGTATGAACA	900
45	CCTGGAAGCA	AAATCATAGT	GCAAAAATAC	ATTTAAGGTG	TGGTCAAAAA	TAAGTCTTTA	960
	ATTGGTAAAT	AATAAGCATT	AATTTTTTAT	AGCCTGTATT	CACAATTCTG	CGGTACCTTA	1020
50	TTGTACCTAA	GGGATTCTAA	AGGTGTTGTC	ACTGTATAAA	ACAGAAAGCA	CTAGGATACA	1080
	AATGAAGCTT	AATTACTAAA	ATGTAATTCT	TGACACTCTT	TCTATAATTA	GCGTTCTTCA	1140
	CCCCACCCC	CACCCCCACC	CCCCTTATTT	TCCTTTTGTC	TCCTGGTGAT	TAGGCCAAAG	1200
55	TCTGGGAGTA	AGGAGAGGAT	TAGGTACTTA	GGAGCAAAGA	AAGAAGTAGC	TIGGAACTIT	1260
	TGAGATGATC	CCTAACATAC	TGTACTACTT	GCTTTTACAA	TGTGTTAGCA	GAAACCAGTG	1320
6 0	GGTTATAATC	TAGAATGATG	TGCTTTCTG	CCAAGTGGTA	ATTCATCTTC	GTTTGCTATG	1380
60							

	TTAAAACTGT AAATACAACA GAACATTAAT AAATATCTCT TGTGTAGCAC CTTTAAAAAA	1440
	AAAAAAAAA AAAAAAAAA AAAAAAAAAA CCCGGGGGG GGCCCCCN	1487
. 5	, .	
	·	
	(2) INFORMATION FOR SEQ ID NO: 99:	
10	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1653 base pairs (B) TYPE: nucleic acid (C) STRANDELNESS: double (D) TOPOLOGY: linear	(
15	·	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 99:	
	GCGACCGCCC CCTTCAGCTA GCTCGCTCGCC TCGCTCTGCT TCCCTGCTGC CGGCTGCGCA	60
20	TGGCTTNGGC GTTGGCGGCG CTGGGGGCGC CTGCGSAGCC GGTACCAGCA	120
	GTTGCAGAAT GAAGAAGAGT CTGGAGAACC TGAACAGGCT GCAGGTGATG CTCCTCCACC	180
25	TTACAGCAGC ATTTCTGCAG AGAGCGCACA TNATTTTGAC TACAAGGATG AGTCTGGGTT	240
	TCCAAAGCCC CCATCTTACA ATGTAGCTAC AACACTGCCC AGTTATGATG AAGCGGAGAG	300
	GACCAAGGCT GAAGCTACTA TCCCTTTGGT TCCTGGGAGA GATGAGGATT TTGTGGGTCG	360
30	GGATGATTTT GATGATGCTG ACCAGCTGAG GATAGGAAAT GATGGGATTT TCATGTTAAC	420
	TPPPPCATG GCATTCCTCT TTAACTGGAT TGGGTTTTTC ÇTGTCTTTTT GCCTGACCAC	480
35	TYCAGCTGCA GGAAGGTATG GGGCCATTTC AGGATTTGGT CTCTCTAA TYAAATGGAT	540
	CCTGATTGTC AGGITTTCCA CCTATTTCCC TGCATTTATG AATTCTCTCT CAAGAAGCAA	600
	GAGAACACCT GCAGGAAGTG AATCAAGATG CAGAACACAG AGGAATAATC ACCTGCTTTA	660
40	AAAAAATAAA GTACTGTTGA AAAGATCATT TCTCTCTATT TGTTCCTAGG TGTAAAATTT	720
	TAATAGITAA TGCAGAATTC TGTAATCATT GAATCATTAG TGGTTAATGT TTGAAAAAGC	780
45	TCTTGCAATC AAGTCTGTGA TGTATTAATA ATGCCTTATA TATTGTTTGT AGTCATTTTA	840
,,,	AGRAGCATGA GCCATGTCCC TGTAGTCGGT AGGGGGCAGT CTTGCTTTAT TCATCCTCCA	900
	TCTCAAAATG AACTTGGAAT TAAATATTGT AAGATATGTA TAATGCTGGC CATTTTAAAG	960
50	GGGTTTTCTC AAAAGTTAAA CTTTTGTTAT GACTGTGTTT TTGCACATAA TCCATATTTG	1020
	CTGTTCAAGT TAATCTAGAA ATTTATTCAA TTCTGTATGA ACACCTGGAA GCAAAATCAT	1080
55	AGTGCAAAAA TACATTTAAG GTGTGGTCAA AAATAAGTCT TTAATTGGTA AATAATAAGC	1140
<i>33</i>	ATTAATTTT TATAGCCTGT ATTCACAATT CTGCGGTACC TTATTGTACC TAAGGGATTC	1200
	TAAAGGTGTT GTCACTGTAT AAAACAGAAA GCACTAGGAT ACAAATGAAG CITAATTACT	1260
60	AAAATGTAAT TCTTGACACT CTTTCTATAA TTAGCGTTCT TCACCCCCAC CCCCACCCCC	1320

	·	
	ACCCCCCTTA TTTTCCTTTT GTCTCCTGGT GATTAGGCCA AAGTCTGGGA GTAAGGAGAG	1380
5	GATTAGGTAC TTAGGAGCAA AGAAAGAAGT AGCTTGGAAC TTTTGAGATG ATCCCTAACA	1440
,	TACTGTACTA CTTGCTTTTA CAATGTGTTA GCAGAAACCA GTGGGTTATA ATGTAGAATG	1500
	ATGTGCTTTC TGCCCAAGTG GTAATTCATC TTGGTTTGCT ATGTTAAAAC TGTAAATACA	. 1560
10	ACAGAACATT AATAAATATC TCTTGTGTAG CACCTTTTAW AAAAAAAAAA AAAAAAAAA	1620
	AAAAAAAAAA AAAAANCCCG GGGGGGGCC CCN	1653
1 5		
15		
	(2) INFORMATION FOR SEQ ID NO: 100:	
	(i) SEQUENCE CHARACTERISTICS:	
20	(A) LENGTH: 1145 base pairs	
	(B) TYPE: nucleic acid (C) STRANDEDNESS: double	
	(D) TOPOLOGY: linear	
25	(wi) CENTERICE DECEPTIONION, CENTER NO. 100.	
2.5	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 100:	
	TTTTTTTTT TTTTTTTT TTGACTGAAC TAAGTGGCTT TTTTATTAGA GAAAGCCAGA	60
30	ATTACAAAAG ACTTCCCTTT TCTTGGGGTA TGGCTGTCTC AGCACAATAC TCAACATAAC	120
	TGCAGAACTG ATGTGGCTCA GGCACCCTGG TTTTAATTCC TTGAGGATCT GGCAATTGGC	180
	TTACGCAAAA GGTCACCATT TGAGGTCCTG CCTTACTAAT TATGTGCTGC CCAACAACTA	240
35	AATTIGIAAT TIGITTTICT CTAGITTGAG CAGGGTCTGA ATTITTTCAT TTATTTCCTT	300
	TTTTGCCAGC AGACAGACTT GAGTCTGTAA AGACAAGCAA ATACACTGAC AGAAGTTTAC	360
40	CATAGTITCT AAAATGTAAA AAAGAAAACC CCCAAAAGAC TCAAGAAAAT TAGACCACAA	420
	ATTITICATI GITCATIGIA GCACTATIGG TAATAAAATA ACAAATGITI GIGCATITITI	480
	ATGTGAAGAT CCTTCTCGTA TTTCATTTGG AAAGATGAGC AAGAGGTCTG CTTCCTTCAT	540
45	TTTACTTCCC CTTCTGTTTT TGAAAGGCAG TTTCGCCAAG CTTAATGCAA GAATATCTGA	600
	CTGTTTAGAA GAAAGATATT GCCACAATCT CTGGATGGTT TTCCAGGGTT GTGTTATTAC	660
50	TGAGCTTCAT CTTTCCAGAA TGAGCAAAAC ACTGTCCAGT CTTTGTTACG ATTTTGTAAT	720
	AAATGTGTAC ATTTTTTTTA AATTTTTGGA CATCACATGA ATAAAGGTAT GTATGTACGA	780
	ATGTGTATAT ATTATATATA TGACATCTAT TTTGGAAAAT GTTTGCCCTG CTGTACCTCA	840
55	TTTTTAGGAG GTGTGCATGG ATGCAATATA TGAAAATGGG ACATTCTGGA ACTGCTGGTC	900
	AGGGGACTTT GTCGCCCTGT GCACTAAAAG GGCCAGATTT TCAGCAGCCA AGGACATCCA	960
	TACCCAACTC AATCTCATCC CACTTAAAAC AACTCAACTC ACACAATTCA CTCTCCCTCT	1020

	TTGAACAGCA GCGTTTCATA GGAAGAGAAA AAAAGATCAA TCTTGTATTT TCTGACCACA	1080
	TAAAGGCTTC TTCTCTTTGT AATAAAGTAG AAAAGCTCTC CTCAAAAAAA AAAAAAAAAA	1140
5	AAAAA	1145
10	(2) INFORMATION FOR SEQ ID NO: 101:	
15	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 734 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 101:	
20	TACCCGGCGG ATTCCAGGAA GGTAAATTTA GTCCTATAAT TTTCAGCTTA ATTATAAACA	60
	AAGGAACAAA TAAGTGGAAG GGCAGCTATT ACCATTCGCT TAGTCAAAAC ATTCGGTTAC	120
25	TGCCCTTTAA TACACTCCTA TCATCAGCAC TTCCACCATG TATTACAAGT CTTGACCCAT	180
	CCCTGTCGTA ACTCCAGTAA AAGTTACJGT TACTAGAAAA TTTTTATCAA TTAACTGACA	240,
	AATAGITICT TITTAAAGTA GITTICTICCA TCTTTATTCT GACTAGCTTC CAAAATGTGT	300
30	TCCCTTTTTG AATCGAGGIT TTTTTGTTTT GTTTTGTTTT CTGAAAAAAT CATACAACTT	360
	TGTGCTTCTA TTGCTTTTTT GTGTTTTGTT AAGCATGTCC CTTGGCCCAA ATGGAAGAGG	420
35	AAATGITTAA TTAATGCTTT TTAGTTTAAA TAAATTGAAT CATTTATAAT AATCAGTGTT	480
	AACAATTTAG TGACCCTTGG TAGGTTAAAG GTTGCATTAT TTATACTTGA GATTTTTTTC	540
	CCCTAACTAT TCTGTTTTTT GTACTTTAAA ACTATGGGGG AAATATCACT GGTCTGTCAA	600
40	GAAACAGCAG TAATTATTAC TGAGTTAAAT TGAAAAGTCC AGTGGACCAG GCATTTCTTA	660
	TATAAATAAA ATTGGTGGTA CTAATGIGAA AAAAAAAAA AAAAAAAACT CGAGGGGGC	720
45	CCGGTACCCT ATTA	734
50	(2) INFORMATION FOR SEQ ID NO: 102: (i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 713 base pairs (B) TYPE: nucleic acid	
55	(C) STRANDEDNESS: double (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 102:	
	CCGCGGGAAC GCTGTCCTGG CTGCCGNCAC CCGAACAGCC TGTCCTGGTG CCCCGGCTCC	60
60		

	CIGCCCCCCC	CCCAGTCATG	ACCCTGCGCC	CCTCACTCCT	CCCGCTCCAT	CTGCTGCTGC	120
	TGCTGCTGCT	CAGTGCGGCG	GIGIGCCCGGG	CTGAGGCTGG	ĢCTCGAAACC	GAAAGTCCCG	180
5	TCCGGACCCT	CCAAGTGGAG	ACCCTGGTGG	AGCCCCCAGA	ACCATGTGCC	GAGCCCGCTG	240
	CTTTTGGAGA	CACGCTTCAC	ATACACTACA	CGGGAAGCTT	GGTAGATGGA	CGTATTATTG	300
10	ACACCTCCCT	GACCAGAGAC	CCTCTGGTTA	TAGAACTTGG	CCAAAAGCAG	GTGATTCCAG	360
10	GTCTGGAGCA	GAGTCTTCTC	GACATGTGTG	TGGGAGAGAA	GCGAAGGGCA	ATCATTCCTT	420
	CTCACTTGGC	CTATGGAAAA	CGGGGATTTC	CACCATCTGT	CCCAGCGGAT	GCAGTGGTGC	480
15	AGTATGACGT	GGAGCTGATT	GCACTAATCC	GAGCCAACTA	CTGGCTAAAG	CTGGTGAAGG	540
	GCATTTTGCC	TCTGGTAGGG	ATGCCCATGG	TCCCACCCTC	CTGGGCCTCA	TTGGGTATCA	600
20	CCTATACAGA	AAGGCCAATA	GACCCAAAGT	CTCCAAAAAG	AAGCTCAAGG	AAGAGAAACG	660
20	AAACAAGAGC	AAAAAGAAAT	AATAAATAAT	АААТТТТААА	AAACTTAAAA	ААА	713
				•			

30

(2) INFORMATION FOR SEQ ID NO: 103:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1080 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 103:

CCGATGTGGA CATCATCCTG TCTATCCCCA TGTTCCTGCG CCTGTACCTG ATCGCCCGAG 60 TCATGCTGCT GCACAGAAGC TCTTCACCGA TGCCTCGTCC CGCAGCATCG GGGCCCTCAA 120 40 CAAGATCAAC TTCAACACCC GCTTTGTCAT GAAGACGCTC ATGACCATCT GCCCTGGCAC 180 TGTGCTGCTC GTGTTCAGCA TCTCTCTGTG GATCATTGCT GCCTGGACCG TCCGTGTCTG 240 TGAAAGTCCT GAATCACCAG CCCAGCCTTC TGGCTCATCA CTTCCTGCTT GGTACCATGA 300 45 CCAGCAGGAC GTAACTAGTA ACTITCTGGG TGCCATGTGG CTCATCTCCA TCACATTCCT 360 TTCCATTGGT TATGGGGACA TGGTGCCCCA CACATACTGT GGGAAAGGTG TCTGTCTCCT 420 50 CACTGGCATC ATGGGTGCAG GCTGCACTGC CCTTGTGGTG GCCGTGGTGG CCCGAAAGCT 480 GGAACTCACC AAAGCGGAGA AGCACGTTCA TAANITCATG ATGGACACTC AGCTCACCAA 540 GCGGATCAAG AATGYTGCAG CCAATGTCCT TSGGGAAACA TGGTTAATCT ATAAACACAC 600 55 AAAGYTGYTA AAGAAGATTG ACCATGCCAA AGTGAGGAAC ACCAGAGGAA GTTCYTCCAA 660 GTATCCACCA GTTGAGGAGC GTCAAGATGG AACAGAGGAA GCTGAGTGAC CAAGCCAACA 720 60 NICTGGTGGA CCTTTCCAAG ATGCAGAATG TCMTGTATGA CTTAATCACA GAACTCAATG 780

	•	
	ACCOGAGCGA AGACCTOGAG AAGCAGATTG GCAGCCTOGA GTCGAAGCTG GAGCATCTCA	840
5	COSCCASCIT CAACTCCCTG COGCTGCTCA TCGCCGACAC CCTGCGCCAG CAGCAGCAGC	.900
3	AGCTCCTGTC TGCCATCATC GAGGCCCGGG GTGTCAGCGT GGCAGTGGGC ACCACCCACA	960
	CCCCAATCTC CGATAGCCCC ATTGGGGTCA GCTCCACCTC CTTCCCGACC CCGTACACAA	1020
10	GTTCAAGCAG TTGCTAAATA AATCTCCCCA CTCCAGAAGC ATTAAAAAAA AAAAAAAAAA	1080
	1	
15	(2) INFORMATION FOR SEQ ID NO: 104:	
20	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 489 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double	
	(D) TOPOLOGY: linear	
05	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 104:	
25	GGCACGAGAG GCTTTGAAGC ATTTTTGTCT GTGCTCCCTG ATCTTCAGGT CACCACCATG	60
	AAGTTCTTAG CAGTCCTGGT ACTCTTGGGA GTTTCCATCT TTCTGGTCTC TGCCCAGAAT	120
30	CCGACAACAG CTGCTCCAGC TGACACGTAT CCAGCTACTG GTCCTGCTGA TGATGAAGCC	180
	CCTGATGCTG AAACCACTGC TGCTGCAACC ACTGCGACCA CTGCTGCTCC TACCACTGCA	240
	ACCACCGCTG CTTCTACCAC TGCTCGTAAA GACATTCCAG TTTTACCCAA ATGGGTTGGG	300
35	GATCTCCCGA ATGGTAGAGT GTGTCCCTGA GATGGAATCA GCTTGAGTCT TCTGCAATTG	360
	GTCACAACTA TTCATGCTTC CTGTGATTTC ATCCAACTAC TTACCTTGCC TACGATATCC	420
40	CCTTTATCTC TAATCAGTTT ATTTTCTTTC AAATAAAAA TAACTATGAG CAACAAAAAA	480
	АААААААА	489
45	(2) INFORMATION FOR SEQ ID NO: 105:	
	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 640 base pairs	•
50	(B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 105:	
55	GCGGTCGCCG CTGTTGTTGT GGTCCCCATG GAGCTGCCGT AGCGGACCCA GCACAGCCAG	60
	GAGCCTCCCG GATGACCTCA GCCCCCCCC ACCACTCCCC GTGGTTGCTG GTGCTCACCT	120
60	TOGRETTING ATGCAATGTT CTTAGGATCC TCCTCCCGTC CTTCTCATCC TTCATGTCCA	180

	GGGTGCTGCA	GAAGGACGCG	GAGCAGGAGT	CACAGATGAG	AGCGGAGATC	CAGGACATGA	240
5	AGCAGGAGCT	CTCCACAGTC	AACATGATGG	ACGAGTTTGC	CAGATATGCC	AGGCTGGAAA	300
J	GAAAGATCAA	CAAGATGACG	GATAAGCTCA	AAACCCATGT	GAAAGCTCGG	ACAGCTCAAT	360
	TAGCCAAGAT	AAAATGGGTG	ATAAGTGTCG	CTTTCTACGT	ATTGCAGGCT	GCCCTGATGA	420
10	TCTCACTCAT	TTGGAAÇTAT	TATTCTGTCC	CTGTGGCTGT	CGTGCCGAGT	AAATGGATAA	480
	CCCTYTAGAC	CGCCTGGTAG	CCTTTCCYAY	TAGAGTAGCA	GCTGCTGTTG	GAATTACTGT	540
15	TOGATTTART	CTGTACAAAT	TGTCCTATTG	TGCTTCACCG	TYCASTGAAC	AGGAGGTGGT	600
	ACAGCCGGAG	TTAAAAACGG	TTTCCNTTCC	AGTTTAAAAT			6 4 0

(2) INFORMATION FOR SEQ ID NO: 106:

(i) SEQUENCE CHARACTERISTICS:

25

30

(A) LENGTH: 1529 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 106:

GGCACNAGA TGGAGCTGCC GTAGCGGACC CAGCACAGCC AGGAGCGTCC GGGATGAGCT 60 CAGCCGCGC CGACCACTGG GCGTGGTTGC TGGTGCTCAG CTTCGTGTTT GGATGCAATG 120 35 TTCTTAGGAT CCTCCCCG TCCTTCTCAT CCTTCATGTC CAGGGTGCTG CAGAAGGACG 180 CGGACAGGAG TCACAGATGA GAGCGGAGAT CCAGGACATG AAGCAGGAGC TCTCCACAGT 240 CAACATGATG GACGAGTITTG CCAGATATGC CAGGCTGGAA AGAAAGATCA ACAAGATGAC 300 40 GGATAAGCTC AAAACCCATG TGAAAGCTCG GACAGCTCAA TTAGCCAAGA TAAAATGGGT 360 GATAAGTGTC GCTTTCTACG TATTGCAGGC TGCCCTGATG ATCTCACTCA TTTGGAAGTA 420 45 TTATTCTGTC CCTGTGGCTG TCGTGCCGAG TAAATGGATA ACCCCTCTAG ACCGCCTGGT 480 AGCCTTTCCT ACTAGAGTAG CAGGTGGTGT TGGAATTACC TGTTGGATTT TAGTCTGTAA 540 CAAAGTTGTC GCTATTGTGC TTCATCCGTT CAGCTGAACA GGAGGATGGA TACAGCCGCG 600 50 AGTAAAAAA CGGATTTCCT CTTCCTAGCT TAAAATCTGA TTTACACTGT TTTGTTTTTT 660 AAGAAACAAA AGTGCATAGT TTAGATTTTT TTTTTGTTGA ATATGTTTGT TCTTGGACTT 720 55 TATGAGATAG TCTTATAAGA ATCACGATTT TCTACACCTG TCATTGAGCC AAGAAAGTCC 780 AGITTATGAC ACGITATGIAC TAGIGAACAC CGICCTCGAT CIGIACGAAA TGIGAAATGI 840 TTAGGGACAT CTCCATGCTG TCACTTGTGA TTTGCCCTCT TATGTATTTT GGTCATATTG 900 60

	CCAACTGGAA	AGTCAAAATT	TTCTAACAAC	TTTAAGTAAG	TTCTTTGAAG	ACTTAGTGCT	960
	GTTTTTAATC	CAGTTTAGAA	AGTAACTTAA	TTTTAATACC	RCTACTAAAA	ATTCGAAAAT	1020
5	TTCTTCTTTA	ATCACATTCA	ATATGGTTAA	AAGAACAACA	CTAATTGACA	TTGCGTGGGC	1080
	TTTTTCTCCC	TTTGTTTAAA	ATGTCATTTG	TTGAGCAAGA	CTTCTATACT	ATTATCTACT	1140
10	TACTTGAGGC	TGTTAATTTT	TCATTACAGT	GTTTTGTAAA	TGTATCCACG	AGACCATGAT	1200
. •	GCATTGTTTT	GİGCTCAACT	TGTGTTTTGT	ATTTAAAGCA	TTTTGAATGA	AGTGTATTTT	1260
	ATAAGCATTT	AATATTTATG	CTCTTTAGAA	TGGAACACAG	AAAACAAACC	TTATAAGTCC	1320
15	TGATTAATCT	GAACCAATAA	CCTGTGTGGC	CTACAAAGTA	TAATTCTATT	AAATGTTCCT	1380
	TAAAACACTT	TTTTCTAATT	AAAATCTTTG	CAAATGCTTG	TGTAACTTCC	TGCCTTACAG	1440
20	CTACTIGITT	GCTGTGAGCC	ACCCGCAACT	GACAAGTGGC	TGTTAACTGA	GTCACCATAT	1500
	CCCAGTAAAG	CTGAATTTTC	TCACTAAAA	ŀ			1529

30

(2) INFORMATION FOR SEQ ID NO: 107:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2435 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 107:

-							
	ATGAAGGGTC	GTTGGTGGGA	AAGATGGCGG	CGACTCTGGG	ACCCCTTGGT	CGTGGCAGCA	60
	GTGGCGRCGA	TGTTTGTCGG	CTCGGGATGG	GTCCAGGATG	TTACTCCTTC	TTCTTTTGTT	120
40	GGGGTCTGGG	CAGGGGCCAC	AGCAAGTCGG	GGCGGGTCAA	ACGTTCGAGT	ACTTGAAACG	180
	GGAGCACTCG	CTGTCGAAGC	CCTACCAGGG	TGTGGGCACA	GGCAGTTCCT	CACTGTGGAA	240
45	TCTGATGGGC	AATGCCATGG	TGATGACCCA	GTATATCCGC	CTTACCCCAG	ATATGCAAAG	300
73	TAAACAGGGT	GCCTTGTGGA	ACCGGGTGCC	ATGTTTCCTG	AGAGACTGGG	AGTTGCAGGT	360
	GCACTTCAAA	ATCCATGGAC	AAGGAAAGAA	GAATCTGCAT	GGGGATGGCT	TGGCAATCTG	420
50	GTACACAAAG	GRWTCGGATG	CAGCCAGGGC	CIGINITIGG	GAAACATGGA	CAAATTTGTG	480
	GGGCTGGGAG	TATTTGTAGA	CACCTACCCC	AATGAGGAGA	AGCAGCAAGA	GCGGGTATTC	540
<i>E E</i>	CCCTRCMTCT	CAGCCATGGT	GAACAACGGC	TCCCTCAGCT	ATGATCATGA	GCGGGATGGG	600
55	CGGCCTACAG	AGCTGGGAGG	CTGCASAGCC	ATTGTCCGCA	ATCTTCATTA	CGACACCTTC	660
	CTGGTGATTC	GCTACGTCAA	GAGGCATTTR	ACGATAATGA	TGGATATTGA	TGGCAAGCAT	720
60	GAGTGGAGGG	ACTGCATTGA	AGTGCCCGGA	GTCCGCCTGC	CCCGCGGCTA	CTACTTCGGC	780

	ACCTCCTCCA	TCACTGGGGA	TCTCTCAGAT	AATCATGATG	TCATTTCCTT	GAAGTTGTTT	840
5	GAACTGACAG	TGGAGAGAAC	CCCAGAAGAG	GAAAAGCTCC	ATCGAGATGT	GITCITGCCC	900
3	TCAGTGGACA	ATATGAAGCT	GCCTGAGATG	ACAGCTCCAC	TGCCGCCCCT	GAGTGGCCTG	960
	GCCCTCTTCC	TCATCGTCTT	TTTCTCCCTG	GGIGITTICT	GTATTTGCCA	TAGTCATTGG	. 1020
10	TATCATACTC	TACAACAAAT	GGCAGGAACA	GAGCCGAAAG	CGCTTCTACT	GAGCCCTCCT	1080
	GCTGCCACCA	CTTTTGTGAC	TGTCACCCAT	GAGGTATGGA	AGGAGCAGGC	ACTGGCCTGA	1140
15	GCATGCAGCC	TGGAGAGTGT	TCTTGTCTCT	AGCAGCTGGT	TGGGGACTAT	ATTCTGTCAC	1200
	TGGAGTTTTG	AATGCAGGGA	CCCCGCATTC	CCATGGTTGT	GCATGGGGAC	ATCTAACTCT	1260
	GGTCTGGGAA	GCCACCCACC	CCAGGCAAT	GCTGCTGTGA	TGTGCCTTTC	CCTGCAGTCC	1320
20	TTCCATGTGG	GAGCAGAGGT	GTGAAGAGAA	TTTACGTGGT	TGTGATGCCA	AAATCACAGA	1380
	ACAGAATTTC	ATAGCCCAGG	CTCCCGTGTT	GTTTGACTCA	GAAGGCCCTT	CTACTTCAGT	1440
25	TTTGAATCCA	CAAAGAATTA	AAAACTGGTA	ACACCACAGG	CTTTCTGACC	ATCCATTCGT	1500
	TGGGTTTTGC	ATTTGACCCA	ACCCTCTGCC	TACCTGAGGA	CCTTTCTTTG	GAAACCAGGA	1560
	TGGAAACTTC	TTCCCTGCCT	TACCTTCCTT	TCACTCCATT	CATTGTCCTC	TCTGTGTGCA	1620
30	ACCTGAGCTG	GGAAAGGCAT	TTGGATGCCT	CTCTGTTGGG	CCTCCCCT	GCAGAACACA	1680
	CCTGCGTTTC	ACTGGCCTTC	ATTAGGŢGGC	CCTAGGGAGA	TĠĢĊŢŢŢĊŢĠ	CTTTGGATCA	1740
35	CTGTTCCCTA	GCATGGGTCT	TGGGTCTATT	GGCATGTCCA	TGGCCTTCCC	AATCAAGTCT	1800
	CTTCAGGCCC	TCAGTGAAGT	TTGGCTAAAG	GTTGGTGTAA	AAATCAAGAG	AAGCCTGGAA	1860
	GACATCATGG	ATGCCATGGA	TTAGCTGTGC	AACTGACCAG	CTCCAGGTTT	GATCAAACCA	1920
40	AAAGCAACAT	TTGTCATGTG	GTCTGACCAT	GTGGAGATGT	TTCTGGACTT	GCTAGAGCCT	1980
	GCTTAGCTGC	ATGTTTTGTA	GTTACGATTT	TTGGAATCCC	ACTITGAGIG	CTGAAAGTGT	2040
45	AAGGAAGCTT	TCTTCTTACA	CCTTGGGCTT	GGATATTGCC	CAGAGAAGAA	ATTTGGCTTT	2100
	TTTTTTNCTT	AATGGACAAG	AGACAGTTGC	TGTTCTCATG	TICCAAGTCT	GAGAGCAACA	2160
	GACCCTCATC	ATCTGTGCCT	GGAAGAGTTC	ACTGTCATTG	AGCAGCACAG	CCTGAGTGCT	2220
50	GCCTCTGTC	AACCCTTATT	CCACTGCCTT	ATTTGACAAG	GGGTTACATG	CTGCTCACCT	2280
	TACTGCCCTG	GGATTAAATC	AGTTACAGGC	CAGAGTCTCC	TTGGAGGCC	TGGAACTCTG	2340
55	AGTCCTCCTA	TGAACCTCTG	TAGCCTAAAT	GAAATTCTTA	AAATCACCGA	TGGAACCAAA	2400
	ААААААААА	ААААААААА	ААААААААА	AAAAN			2435

	(2) INFORMATION FOR SEQ ID NO: 108:	
5	(i) SEQUENCE CHARACTERISTICS: (A) LEWGTH: 805 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
10	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 108:	
	ATGAAACTTA AGAATTGAAT TGGAAAGACT TCTCAAAGAG AATTGTATGT AACGATGTTG	60
	TATTGATTT TAAGAAAGTA ATTTAATTTG TAAAACTTCT GCTCGTTTAC ACTGCACATT	120
15	GAATACAGGT AACTAATTGG AAGGAGAGGG GAGGTCACTC TITTGATGGT GGCCCTGAAC	180
	CTCATTCTGG TTCCCTGCTG CGCTGCTTGG TGTGACCCAC GGAGGATCCA CTCCCAGGAT	240
20	GACGTGCTCC GTAGCTCTGC TGCTGATACT GGGTCTGCGA TGCAGCGGCG TGAGGCCTGG	300
-0	GCTGGTTGGA GAAGGTCACA ACCCTTCTCT GTTGGTCTGC CTTCTGCTGA AAGACTCGAG	360
	AACCAACCAG GGAAGCTGTC CTGGAGGTCC CTGGTCGGAG AGGGACATAG AATCTGTGAC	420
25	CTCTGACAAC TGTGAAGCCA CCCTGGGCTA CAGAAACCAC AGTCTTCCCA GCAATTATTA	480
	CAATTCTTGA ATTCCTTGGG GATTTTTTAC TGCCCTTTCA AAGCACTTAA GTGTTAGATC	5 40
30	TAACGTGTTC CAGTGTCTGT CTGAGGTGAC TTAAAAAATC AGAACAAAAC TTCTATTATC	600
,,	CAGAGTCATG GGAGAGTACA CCCTTTCCAG GAATAATGTT TTGGGAAACA CTGAAATGAA	660
	ATCTTCCCAG TATTATAAAT TGTGTATTTA AAAAAAAGAA ACTTTTCTGA ATGCCTACTG	720
35	GCGGTGTATA CCAGGCAGTG TGCCAGTTTA AAAAGATGAA AAAGAATAAA AACTTTTGAG	780
	GAACAAAAA AAAAAAAAA AAATT	805
ю		
	(2) INFORMATION FOR SEQ ID NO: 109:	
15	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1166 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
50	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 109:	
	GGCACGAGAG GCGCCAGTCG CAGGTGTGCT GCTGAGGCGT GAGAATGGCG TCCCGCGGCC	60
55	GGCGTCCGGA GCATGGCGGA CCCCCAGAGC TGTTTTATGA CGAGACAGAA GCCCGGAAAT	120
,,,	ACCITCGCAA CTCACGGATG ATTGATATCC AGACCAGGAT GGCTGGGCGA GCATTGGAGC	180
	TTCTTTATCT GCCAGAGAAT AAGCCCTGTT ACCTGCTGGA TATTGGCTGT GGCACTGGGC	240

TGAGTGGAAG TTATCTGTCA GATGAAGGGC ACTATTGGGT GGGCCTGGAT ATCAGCCCTG

	CCATGCTGGA	TGAGGCTGTG	GACCGAGAGA	TAGAGGGAGA	CCTGCTGCTG	GGGGATATGG	· 360
5	GCCAGGGCAT	CCCATTCAAG	CCAGGCACAT	TTGATGGTTG	CATCAGCATT	TCTGCTGTGC	420
3	AGTGGCTCTG	TAATGCTAAC	AAGAAGTCTG	AAAACCCTGC	CAAGCGCCTG	TACTGCTTTT	480
	TIGCTICICT	TTTTTCTGTT	CTCGTCCGGG	GATCCCGAGC	TGTCCTGCAG	CTGTACCCTG	540
10	AGAACTCAGA	GCAGTTGGAG	CTGATCACAA	CCCAGGCCAC	AAAGGCAGGC	TICICCGGIG	600
	GCATGGTGGT	AGACTACCCT	AACAGTGCCA	AAGCAAAGAA	ATTCTACCTC	TGCTTGTTTT	660
15	CTGGGCCTTC	GACCTTTATA	CCAGAGGGGC	TGAGTGAAAA	TĊAGGATGAA	GTTGAACCCA	720
13	GGGAGTCTGT	GTTCACCAAT	GAGAGGTTCC	CATTAAGGAT	GTCGAGGCGG	GGAATGGTGA	780
	GGAAGAGTCG	GGCATGGGTG	CTGGAGAAGA	AGGAGCGGCA	CAGGCGCCAG	GGCAGGGAAG	840
20	TCAGACCTGA	CACCCAGTAC	ACCGGCCGCA	AGCGCAAGCC	CCGCTTCTAA	GTCACCACGC	900
	GGTTCTGGAA	AGGCACTIGC	CTCTGCACTT	TTCTATATTG	TTCAGCTGAC	AAAGTAGTAT	960
25	TTTAGAAAAG	TTCTAAAGTT	ATAAAAATGT	TTTCTGCAGT	AAAAAAAAAG	TTCTCTGGGC	1020
23	CGGCCTGCT	GGCTCACANC	TGTAATCCCA	GCACCTTGGG	AGGCTGAGGT	GGGAGGATCA	1080
	TTTGAGGCCA	GGAGTTTGAG	ACCTGCCTGG	GCAACATAAT	GAAACTTCCT	TTCCAGGGAG	1140
30	АААААААА	АААААААА	ACTCGA				1166

35 (2) INFORMATION FOR SEQ ID NO: 110:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 586 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 110:

45	AGAGCGGACG	AAGCTGGATA	ACAGGGGACC	GATGATGTGG	CGACCATCAG	TTCTGCTGCT	60
	TCTGTTGCTA	CTGAGGCACG	GGGCCCAGGG	GAAGCCATCC	CCAGACGCAG	GCCCTCATGG	120
50	CCAGGGGAGG	GTGCACCAGG	CCCCCCT	GAGCGACGCT	CCCCATGATG	ACGCCCACGG	180
30	GAACTTCCAG	TACGACCATG	AGGCTTTCCT	GGGACGGGAA	GTGGCCAAGG	AATTCGACCA	240
	ACTCACCCCA	GAGGAAAGCC	AGGCCCGTCT	GGGGGGATC	GTGGACCGCA	TGGACCGCGC	300
55	GGGGGACGGC	GACGGCTGGG	TGTCGCTGGC	CGAGCTTCGC	GCGTGGATCG	CGCACACGCA	360
	GCAGCGGCAC	ATACGGGACT	CGGTGAGCGC	GGCCTGGGAC	ACGTACGACA	CGGACCGCGA	420
	CCCCCCTCTC	GGTTGGGAGG	AGCTGCGCAA	CGYCACCTAT	GCCACTASG	SGCCCGKTGA	480

60

	AGAATTICAT GACGIOGAGG AIGCAGAGAC YTACAAAAAG ATGCTGGYTC GGGACG	AGCG 540
·5	GCGTTTCCGG GTGGCCGACC AGGATGGGGA CTCGATGGCC ACTCGA	586
3		
	(2) INFORMATION FOR SEQ ID NO: 111:	
10	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1134 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
15	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 111:	
20	ACCCATTGAG CAGAAGGAGG CCAGGTGGGA AAGCTCCTGG GAAGAGCAGC CAGACTY	
20	ACTGGGCTGC TTGAGTCCTG AGTCACAATT CAGAATTCCT GGGCTCCCTG GGTGCAT	
	ATCATTCCAG TIGAAAGTIT GCTTCCTTCC AGTCATGIGG CTCTTCATTC TACTCTC	TTT 180
25	GGCTCTCATT TCAGATGCCA TGGTCATGGA TGAAAAGGTC AAGAGAAGTT TGTGCTC	GAC 240
	ACGCCTTCTG CCATCTGCAA CTACAATGCC CAYTACAAGA ATCACCCCAA ATACTGC	FTGC 300
	CGAGGYTATT TCCGTGAYTA CTGCAACATC ATCGCCTTCT CCCCTAACAG CACCAAT	CAT 360
30	GTGGCCCTGA AGGACACAGG GAACCAGCTC ATTGTCACTA TGTCCTGCCT GAACAAA	NAA 420
	GACACGGGCT GGTACTGGTG TGGCATCCAR CGGGACTTTG CMAGGGATGA CATGGAT	TTT 480
35	ACAGAGCTGA TTGTAACTGA CGACAAAGGA ACCCTGGCCA ATGACTTTTG GTCTGGG	AAA 540
<i>J J</i>	GACCTATCAG GCAACAAAAC CAGAAGCTGC AAGGCTCCCA AAGTTGTCCG CAAGCTG	ACC 600
	GCTCCAGGAC GTCCATTCTC ATCATTTGCA TACTGATCAC GGGTTTGGGA ATCATCT	CTG 660
40	TAATCAGTCA TITGACCAAA AGGAGGAGAA GTCAAAGGAA TAGAAGGGTA GGCAACA	CTT 720
	TGAAGCCCTT CTCGCGTGTC CTGACTCCAA AGGAAATGGC TCCTACTGAA CAGATGT	
	TGAAGWITTT TITAATTTAG TINCATAAAG TGATGNCTAC AACAGAWIAA TCACCCA	. = •
45	CAACTGGCCC CACACCTCAG AGACTGATTC TGATCTCCCA GGAATTCTGA AGGACCC	
	ATCCTTGACA ACAATCATTT GCAGCCAGGT AGCAACGGCR GTAGTCAGAG GAGCTAT	
50	AGACCACACC CAAGCAAGGC TGCCCTCAAA TAACATCTCA AGATCTTAGT TCTTATG	
	TCCATCAGTC AGAAGTGAAG AAGAGGTGGA GAATCTKGAT TGGGGACCAG GAAATCA	
		CTT 1080
55	GTATTITGTT AGCCAATAAA TTCCTAGCCA GTGTTGAATG AAAAAAAAAA	1134

⁽²⁾ INFORMATION FOR SEQ ID NO: 112:

1:1	CLEAST LEAVINGE	CHARACTERISTICS
(11	SEUUENCE	CUMMACIENTOITES

(A) LENGTH: 1333 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 112:

10	CACTTTAAAG	CTCTGCTGAG	GGAGTTCGGA	GCCCAGGCTT	TCAGGCGACC	TCTGCCCTCC	60
10	CTCCCTCTCC	TCACCCTCCC	TCTCTTCCTG	CAGGGCCTGG	GAAGGGCTTT	GAGGGAGCCT	120
	GGGAGCCATG	TGAAGAGGGG	CACGCCTGGG	CTCTCCCACA	GTTTAGATCC	AGTTGGAGGT	180
15	TCTCCCTGGC	TCCTGCAGGC	CTGCGGGGAT	CTCTCCCCAC	TTCAGGCCTC	CGGCAGCTGC	240
	CTGCCCTCTT	GTCTGTGCTT	CAGCCCTGCA	CAAAAGCAGC	TTGGTGACAC	CACTCAGCCA	300
20	CCCAGAGTAC	GTGTTTACAG	GCTTTCCAGA	TCACCTTCCT	GTGGGGTGAA	CGTAATGAGG	360
20	CGGGGCTGGT	CCTTGGAATT	TCCCCTGGAA	AATGGTAACA	GACTCCATCC	TTGACCCCGG	420
	GATGAGCATG	AAGGCATTGT	CCCAAAGGCA	GAGGCCACCG	TGGTAGGAAT	TCCACCAAGG	480
25	CCAGAAGGGA	AAAAGGAAGA	ACCCACCGTG	TCTGGCTGTG	CGGGCCCTGG	GGAGGGTCGT	540
	GAGTGCAGCC	CCTCTCTACT	TCYGTGCCTT	TGTAAAACGT	GTAGATAACC	GCAGTGGTTG	600
30	GCTGAGCCAA	GAACTCTCCT	AAATCAGTGG	CTTTCTCCCC	ACCCCTTGCT	GGGGAGTCAT	660
50	AAAAAATTTT	ATCTGTGGGA	TATAAAATTG	GCCTCCTGCT	GCTTCAGCCT	ACCTCTCCCT	720
	CIGCIGACTI	AATGTCGTGA	TTCTGTTTCT	TCAGATATTT	AAGGCTGTTA	GGTTGTGTGA	780
35	GCCTTGAAGT	GIGIGIGIGI	GTCCCAGCGA	CTGTCCACTG	TCCAGGAGAT	GCATGTCTTT	840
	GTATTGGAGA	TATTTCTGTA	ACTCATTCTC	TTGGTGCTCA	CGATTGCCAT	GGCCATAGGG	900
40	CCACAGTGCC	GTATCTGCTG	CAGACATGAT	TGTTTCTTGT	TCTAGAGGTT	TTCTTGTTTT	960
10	CGAATCTTGC	CTGATGAATC	CAGCCAGACC	AAGGGGCCTA	GATTTGACCT	CIGICCIGGG	1020
	CTCCTGGGCC	AGGTGCAGGA	ACATCTGAGG	CCACTCTGCT	GGCCACCTCC	AGTGGGTGCT	1080
45	GACCACAGGA	TGGGCTTTGT	TTACACTCAT	TTTCACCCTG	ATTCTTGCCC	CCACTTTCAT	1140
	AAAAGAAACT	TCAAAATGCT	GACGCTTTGG	AGAGTAAGAA	AATCAATCTT	GGCTGGGCAC	1200
50	GGIGGCTCCT	GCCTGTGATC	CTAGCACTTT	GGGAGGCTGA	AGCTGAAGGA	TCACTTGAGC	1260
50	TCAGGAGTTG	GAGACCAACC	CTGGCAACAT	AACAAGACCC	TGTCTCTACA	АААААААА	1320
	AAAAAAAACT	CGA					1333

55

(2) INFORMATION FOR SEQ ID NO: 113:

60 (i) SEQUENCE CHARACTERISTICS:

	(A) LENGTH: 1015 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double	
· 5	(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 113:	
	GGCACGAGCG GCACGAGCGG CACGAGGTGA CTTCAAGTGT CGGATCTTTT CAGCCTACAT	. 60
10	CAAGGAGGTG GAGGAACGGC CGGCACCCAC CCCGTGGGCT CCAAGATGCC CTTTGGGGAA	120
	CTGATGTTCG AATCCAGCAG TAGCTGCGGC TGGGTACATG GCGTCTGTTT CTCAGCCAGC	180
15	GGGAGCCGCG TGGCCTGGGT AAGCCACGAC AGCACCGTCT GCCTGGCTGA TGCCGACAAG	240
	AAGATGGCCG TCGCGACTCT GGCCTCTGAA ACACTACCAC TGCTGGCGCT GACCTTCATC	300
	ACAGACAACA GCCTGGTGGC AGCGGGCCAC GACTGCTTCC CGGTGCTGTT CACCTATGAC	360
20	GCCGCCGCGG GGATGCTGAG CTTCGGCGGG CGGCTGGACG TTCCTAAGCA GAGCTCGCAG	420
	CGTGGCTTGA CGGCCCGCGA GCGCTTCCAG AACCTGGACA AGAAGGCGAG CTCCGAGGGT	480
25	GGCACGGCTG CGGGCGCGGG CCTAGACTCG CTGCACAAGA ACAGCGTCAG CCAGATCTCG	540
	GTGCTCAGCG GCGCCAAGGC CAAGTGCTCG CAGTTCTGCA CCACTGGCAT GGATGGCGC	600
	ATGAGTATCT GGGATGTGAA GAGCTTGGAG TCAGCCTTGA AGGACCTCAA GATCAAATGA	660
30	CCTGTGAGGA ATATGTTGCC TTCATCCTAG CTGCTGGGGA AGCGGGGAGA GGGGTCAGGG	720
	AGGCTAATGG TTGCTTTGCT GAATGTTTCT GGGGTACCAA TACGAGTTCC CATAGGGGCT	780
35	GCTCCCTCAA AAAGGGAGGG GACAGATGGG GAGCTTTTCT TACCTATTCA AGGAATACGT	840
	GCCTTTTTCT TAAATGCTTT CATTTATTGA AAAAAAAAA AAATGCCCCC AAAGCACTAT	900
40	GCTGGTCATG AACTGCTTCA AAATGTGGAG GTAATAAAAT GCAACTGTGT AAAAAAAAA	960
40	AAAAAAAAA AAAAAAAAA AAAAAAAAA AAAAAAAA	1015
45	(2) INFORMATION FOR SEQ ID NO: 114:	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 1076 base pairs	
50	(B) TYPE: nucleic acid	
50	(C) STRANDEDNESS: double (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 114:	
55	GGCACGAGGG GAAAGCCATG CTCCCAGGAC TCCTTCCTTG CAGCCTTAAA TCGGTCTGTA	60
	CGGAAAATTC CGCGCCTTAG AAACCCACGC TTGGGTGTAA CTTATTATTG TTCTTCCTGA	120
60	CCTACTTCCT GTTTATCACT TCCGGGTTCA TCATTTTGGC ATTTCGGTGA TCGGGTTGGA	180

	ACTATTGAAG	CCCCCTTTCA	GGTTCTTTTC	CCCATTTTCC	CTTTGAAAGG	AAGACTTCTG	240
	GCTTCTCCTA A	AATCTCCGTT	CTCTGGGTAA	GGGGAGTCCA	AGCCTCTGTC	ATGAGGAACG	300
5	GAAATGCGAG (GCCTCGGGT	GTTACTCTAA	AATCCGCCCT	CAGCTTGCAC	GCCGGAAGCT	360
	GCGATTCCTG (CAGCGGAAGA	GGCGTGATCT	GCCTTCGAC	TCGCTATGTC	CACTAACAAT	420
10	ATGTCGGACC (CACGGAGGCC	GAACAAAGTG	CTGAGGTACA	AGCCCCCCCC	GAGCGAATGT	48,0
10	AACCCGGCCT T	rggacgaccc	GACGCCGGAC	TACATGAACC	TGCTGGGCAT	GATCTTCAGC	540
	ATGTGCGGCC T	rcatgcitaa	GCTGAAGTGG	TCTCCTTCCC	TCGCTGTCTA	CTCCTCCTTC	600
15	ATCAGCTTTG (CCAACTCTCG	GAGCTCGGAG	GACACGAAGC	AAATGATGAG	TAGCTTCATG	660
	CTGTCCATCT (CTGCCGTGGT	GATGTCCTAT	CTGCAGAATC	CTCAGCCCAT	GACGCCCCCA	720
20	TGGTGATACC A	AGCCTAGAAG	GCTCACATTT	TGGACCCTGT	CTATCCACTA	GCCTGGGCT	780
20	TTGGCTGCTA A	AACCIGCIGC	CTTCAGCTGC	CATCCTGGAC	TTCCCTGAAT	GAGGCCGTCT	840
	CGGTGCCCCC 1	AGCTGGATAG	AGGGAACCTG	GCCCTTTCCT	AGGGAACACC	CTAGGCTTAC	900
25 .	CCCTCCTGCC 1	recertecee	TGCCTGCTGC	TGGGGGAGAT	GCTGTCCATG	TTTCTAGGGG	960
	TATTCATTTG (chreregar	GAAACCTGTT	GTTAATAAAG	TTTTTCACTC	TGAAAAAAA	1020
30	AAAAAAAA I	RAAAACNCGN	GGGGGGGCCC	GGAACCCAAT	TCSCCGGATA	GTGAGT	1076
					1		
	(2) INFORMAT	TION FOR SE	20 TD NO: 11	۱۲۰	1 1		
35	(2) INFORMAT				1 1		
	_	SEQUENCE CH (A) LENK	HARACTERIST FTH: 1487 b	ICS: ase pairs	1 1		
35	_	SEQUENCE CH (A) LENC (B) TYPI (C) STRI	HARACTERIST STH: 1487 b E: nucleic ANDEDNESS:	ICS: ase pairs acid double	1 1	·	
	(i) :	SEQUENCE CH (A) LENC (B) TYPI (C) STRI (D) TOPO	HARACTERIST: FTH: 1487 b E: nucleic ANDEDNESS: DLOGY: line	ICS: ase pairs acid double ar	, '		
35	(i) ((xi)	SEQUENCE CH (A) LENK (B) TYPI (C) STRA (D) TOPO SEQUENCE I	HARACTERIST: FIH: 1487 b E: nucleic ANDEDNESS: DLOGY: line DESCRIPTION	ICS: ase pairs acid double ar : SEQ ID NO	: 115:		
35	(i) (xi) CCGCTGCTGA T	SEQUENCE CH (A) LENG (B) TYPI (C) STRI (D) TOPO SEQUENCE I	HARACTERIST: FIH: 1487 b E: nucleic ANDEDNESS: DLOGY: line DESCRIPTION ATCCCCCGGG	ICS: ase pairs acid double ar : SEQ ID NO	: 115:		60
35 40	(xi) CCGCTGCTGA T	SEQUENCE CH (A) LENG (B) TYPI (C) STRI (D) TOPO SEQUENCE I FAACTATGGC	HARACTERIST: FIH: 1487 b E: nucleic ANDEDNESS: DLOGY: line DESCRIPTION ATCCCCCGGG CCTGCAGGCT	ICS: ase pairs acid double ar : SEQ ID NO CCTGCAGGAA CCTCGCGGGT	: 115: TTCGGCACGG GGAGCCCACC	CAAGACATCA	
35 40 45	(i) (xi) CCGCTGCTGA T	SEQUENCE CH (A) LENG (B) TYPI (C) STRI (D) TOPO SEQUENCE I FAACTATGGC	HARACTERIST: FIH: 1487 b E: nucleic ANDEDNESS: DLOGY: line DESCRIPTION ATCCCCCGGG CCTGCAGGCT	ICS: ase pairs acid double ar : SEQ ID NO CCTGCAGGAA CCTCGCGGGT	: 115: TTCGGCACGG GGAGCCCACC	CAAGACATCA	
35 40	(xi) CCGCTGCTGA T	SEQUENCE CH (A) LENK (B) TYPH (C) STRA (D) TOPK SEQUENCE INTERPORT OF THE CONTROL OF THE CO	HARACTERIST: TH: 1487 b E: nucleic ANDEDNESS: DLOGY: line DESCRIPTION ATCCCCCGGG CCTGCAGGCT GGCCAGGACG	ICS: ase pairs acid double ar : SEQ ID NO CCTGCAGGAA CGTCGCGGGT	: 115: TTCGGCACGG GGAGCCCACC CCGGAACCTG	CAAGACATCA	120 180
35 40 45	(xi) CCGCTGCTGA T CCGCCTGGCT C GCATCAGCGA C	SEQUENCE CHECK (A) LENK (B) TYPH (C) STRUM (D) TOPK SEQUENCE INTRACTATOGIC COTTOCTOR (C) CCTGCTGNCA CCAGCTGGGG	HARACTERIST: FIH: 1487 b E: nucleic ANDEDNESS: DLOGY: line DESCRIPTION ATCCCCCGGG CCTGCAGGCT GGCCAGGACG	ICS: ase pairs acid double ar : SEQ ID NO CCTGCAGGAA CGTCGCGGGT TGCCCGTGTT	: 115: TTCGGCACGG GGAGCCCACC CCGGAACCTG GGGCACCCGG	CAAGACATCA TCCCTGCTGG GAGAGGCGCCC	120 180
35 40 45	(xi) (xi) CCGCTGCTGA T CCGCCTGGCT C GCATCAGCGA C TGGTGGGTGT C	SEQUENCE CHECK (A) LENK (B) TYPH (C) STRUCT (D) TOPK SEQUENCE IN TAACTATGGC COTGCTGNCA CCAGCTGGGG	HARACTERIST: FIH: 1487 b E: nucleic ANDEDNESS: DLOGY: line DESCRIPTION ATCCCCCGGG CCTGCAGGCT GGCCAGGACG TTCTCACTGC GGCGAGCACA	ICS: ase pairs acid double ar : SEQ ID NO CCTGCAGGAA CGTCGCGGGT TGCCCCGTGTT TATTCCACCT	: 115: TTCGGCACGG GGAGCCCACC CCGGAACCTG GGGCACCCGG	CAAGACATCA TCCCTGCTGG GAGAGGCGCCC ACGGCCCAGC	120 180 240 300
35 40 45 50	(xi) CCGCTGCTGA T CCGCCTGGCT C GCATCAGCGA C TGGTGGGTGT C GGCCGCATGC C	SEQUENCE CHECK (A) LENG (B) TYPH (C) STRUCT (D) TOPK SEQUENCE IN TAACTATGGC COTGCTGNCA CCAGCTGGGG CGGCCCGTG GGASGAGCCA	HARACTERIST: FIH: 1487 b E: nucleic ANDEDNESS: DLOGY: line DESCRIPTION ATCCCCCGGG CCTGCAGGCT GGCCAGGACG TTCTCACTGC GGCGAGCACA TGGCTCCGGG	ICS: ase pairs acid double ar : SEQ ID NO CCTGCAGGAA CGTCGCGGGT TGCCCGTGTT TATTCCACCT CCCCCCTGTT	: 115: TTCGGCACGG GGAGCCCACC CCGGAACCTG GGGCACCCGG GGCCCCTGCC CTACCAGGTG	CAAGACATCA TCCCTGCTGG GAGAGGCGCCC ACGGCCCCAGC GGCATACTGT	120 180 240 300
35 40 45 50	(xi) CCGCTGCTGA T CCGCCTGGCT C GCATCAGCGA C TGGTGGGTGT C GGCCGCATGC C CCCTGCTGCT C	SEQUENCE CHECK (A) LENK (B) TYPH (C) STRUCK (D) TOPK SEQUENCE IN TARCTATGGC COTTENCA COAGCTGGGG CGCCCGTG SGASGAGCCA CTGGAAGCAC CAGGCTCATC	HARACTERIST: FIH: 1487 b E: nucleic ANDEDNESS: DLOGY: line DESCRIPTION ATCCCCCGGG CCTGCAGGCT GGCCAGGACG TTCTCACTGC GGCGAGCACA TGGCTCCGGG	ICS: ase pairs acid double ar : SEQ ID NO CCTGCAGGAA CGTCGCGGGT TGCCCGTGTT TATTCCACCT CCCCCCTGTT AGCSGGCTTT CCCAGACCTA	: 115: TTCGGCACGG GGAGCCCACC CCGGAACCTG GGGCACCCGG GGCCCCTGCC CTACCAGGTG CATGGCCATG	CAAGACATCA TCCCTGCTGG GAGAGGCGCC ACGGCCCAGC GGCATACTGT TACCTCACCT	120 180 240 300 360

	ACTICICAGG CCICCIGGIG ATCCIGGCCT TIGCCGCCTG GGTGGCGCTG GCGGAGGGAC	600
5	TGGGTGTGGC CGTGTACGCA GCGGCTGTGC TGCTGGGTGC TGGCTGTGCC ACCATCCTCG	. 660
3	TCACCTCGCT GGCCATGACG GCCGACCTCA TCGGTCCCCA CACGAACAGC GGAGCKTTCG	720
	TGTACGCTC CATGAGCTTC TTGGATAAGG TGGCCAATGG GCTGGCAGTC ATGGCCATCC	780
10	AGAGCCTGCA CCCTTGCCCC TCAGAGCTCT GCTGCAGGCC CTGCGTGAGC TTTTACCACT	840
	GGGCGATGGT GGCTGTGACG GGCGGCGTGG GCGTGGCCGC TGCCCTGTGT CTCTGTAGCC	900
15	TCCTGCTGTG GCCGACCCGC CTGCGACGCT GATGAGACCT GCACGCANTG GCTCACAGCA	960
15	GCACGATTTG TGACAGCCCG AGGCGGAGAA CACCGAACAC CCAGTGAAGG TGAGGGGATC	1020
	AGCACGGCGC GGCCACCCAC GCACCCACGC GCTGGAATGA GACTCAGCCA CAAGGAGGTG	1080
20	CGAAGCTCTG ACCCAGGCCA CAGTGCGGAT GCACCTTGAG GATGTCACGC TCAGTGAGAG	1140
	ACACCAGACA CAGAAGGGTA CGCTGTGATC CCACTTCTAT GAAATGTCCA GGACAGACCA	1200
25	ATCCACAGAA TCAGGGAGAG GATTCGTGGG TGCCGGGACT GGGGAGGGGG ACCTGGGGGT	1260
23	GACTAGGTGA CATAATGGGG ACAGGGCTGC CTTCTGGGTG ATGAGAATGT TCTGGAATCA	1320
	GATGGGATGG CTGCACGGCG TGGTGAAGGT ACTGAACGCC ACCTCACTGT AAGACGGTAG	1380
30	ATTTTGTATT TTACCACAAT AAACAAAACA AAACAAAACC AAAAAAAA	1440
	AAAAAAAAGG AATTCGATAT CAAGCTTATC GATACCGTCG ACCTCGA	1487
35	į į	
33	(2) THEODIVINION FOR ONE TO NO. 116	
	(2) INFORMATION FOR SEQ ID NO: 116:	
40	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1350 base pairs	
	(B) TYPE: nucleic acid (C) STRANDEDNESS: double	
45	(D) TOPOLOGY: linear	
43	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 116:	
	GGCACGAGTG CGCANGCGTG GGGCTCTCTC CTTGTCAGTC GGCGCCGCGT GCGGGCTGGT	60
50	GGCTCTGTGG CAGCGGCGGC GGCAGGACTC CGGCACTATG AGCGGCTTCA GCACCGAGGA	120
	GCGCGCCGCG CCTTCTCCCT GGAGTACCGA GTCTTCCTCA AAAATGAGAA AGGACAATAT	
55	ATATCTCCAT TTCATGATAT TCCAATTTAT GCAGATAAGG ATGTGTTTCA CATGGTAGTT	240
<i>JJ</i>	GAAGTACCAC GCTGGTCTAA TGCAAAAATG GAGATTGCTA CAAAGGACCC TTTAAACCCT	
	ATTAAACAAG ATGTGAAAAA AGGAAAACTT CGCTATGTTG CGAATTTGTT CCCGTATAAA	360
60	GGATATATCT GGAACTATGG TGCCATCCCT CAGACTTGGG AAGACCCAGG GCACAATGAT	420

360

420

480

	AAACATACTG GCTGTTGTGG TGACAATGAC CCAATTGATG TGTGTGAAAT TGGAAGCAAG	480
	GTATGTGCAA GAGGTGAAAT AATTGGCGTG AAAGTTCTAG GCATATTGGC TATGATTGAC	540
5	GAAGGGAAA CCGACTGGAA AGTCATTGCC ATTAATGTGG ATGATCCTGA TGCAGCCAAT	600
	TATAATGATA TCAATGATGT CAAACGGCTG AAACCTGGCT ACTTAGAAGC TACTGTGGAC	660
10	TGGTTTAGAA GGTATAAGGT TCCTGATGGA AAACCAGAAA ATGAGTTTGC GTTTAATGCA	720
10	GAATTTAAAG ATAAGGACTT TGCCATTGAT ATTATTAAAA GCACTCATGA CCATTGGAAA	780
	GCATTAGTGA CTAAGAAAAC GAATGGAAAA GGAATCAGTT GCATGAATAC AACTTTGTCT	840
15 .	GAGAGCCCCT TCAAGTGTGA TCCTGATGCT GCCAGAGCCA TTGTGGATGC TTTACCACCA	900
	CCCTGTGAAT CTGCCTGCAC AGTACCAACA GACGTGGATA AGTGGTTCCA TCACCAGAAA	960
20	AACTAATGAG ATTTCTCTGG AATACAAGCT GATATTGCTA CATCGTGTTC ATCTGGATGT	1020
20	ATTAGAAGTA AAAGTAGTAG CTTTTCAAAG CTTTAAATTT GTAGAACTCA TCTAACTAAA	1080
	GTAAATTCTG CTGTGACTAA TCCAATATAC TCAGAATGTT ATCCATCTAA AGCATTTTTC	1140
25	ATATCTCAAC TAAGATAACT TTTAGCACAT GCTTAAATAT CAAAGCAGTT GTCATTTGGA	1200
	AGTCACTTGT GAATAGATGT GCAAGGGGAG CACATATTGG ATGTATATGT TACCATATGT	1260
20	TAGGAAATAA AATTATTTTG CTGAAAAAAA AAAAAAAAAA	1320
30	CCCCATTIGG CCCTTIGGGG GGNGGTTTTA	1350
35	'	
	(2) INFORMATION FOR SEQ ID NO: 117:	
	(i) SEQUENCE CHARACTERISTICS:	
40	(A) LENGTH: 2527 base pairs	
-10	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double (D) TOPOLOGY: linear	
45	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 117:	
7.5	CTCTTGCTAC CTTCCCGGCG CAGAGAACCC CGGCTGCTCA GCGCGCTCCG GGGTCATGGA	60
	GATCCCCGGG AGCCTGTGCA AGAAAGTCAA GCTGAGCAAT AACGCGCAGA ACTGGGGAAT	120
50	GCAGAGAGCA ACCAATGTCA CCTACCAAGC CCATCATGTC AGCAGGAACA AGAGAGGTCA	180
	GGTGGTGGG ACCAGAGGTG GCTTTCGTGG TTGCACAGTT TGGCTAACAG GCTTGTCTGG	240

AGCGGGAAAG ACTACTGTGA GCATGGCCTT GGAGGAGTAC CTGGTTTGTC ATGGTATTCC

ATGCTACACT CTGGATGGTG ACAATATTCG TCAAGGTCTC AATAAAAATC TTGGCTTTAG

TCCTGAAGAC AGAGAAGAGA ATGTTCGACG CATCGCAGAA GTTGCTAAAC TGTTTGCAGA

TGCTGGCTTA GTGTGCATCA CAAGTTTCAT ATCACCTTAC ACTCAGGATC GCAACAATGC

55

	AAGGCAAATT CATGAAGGTG CAAGTTTACC GTTTTTTGAA GTATTTGTTG ATGCTCCTCT	540
. 5	GCATGTTTGT GAACAGAGGG ATGTCAAAGG ACTCTACAAA AAAGCCCGGG CAGGAGAAAT	600
3	TAAAGGTTTC ACTGGGATCG ATTCTGAATA TGAAAAGCCA GAGGCCCCTG AGTTGGTGCT	660
	GAAAACAGAC TCCTGTGATG TAAATGACTG TGTCCAGCAA GTTGTGGAAC TTCTACAGGA	720
10	ACGGGATATT GTACCTGTGG ATGCATCTTA TGAAGTAAAA GAACTATATG TGCCAGAAAA	780
	TAAACTTCAT TTGGCAAAAA CAGATGCGGA AACATTACCA GCACTGAAAA TTAATAAAGT	840
15	GGATATGCAG TGGGTGCAGG TTTTTGGCAGA AGGTTGGGCA ACCCCATTGA ATGGCTTTAT	900
13	GAGAGAGAG GAGTACTTGC AGTGCCTTCA TTTTGATTGT CTTCTGGATG GAGGTGTCAT	960
	TAACTTGTCA GTACCTATAG TTCTGACTGC GACTCATGAA GATAAAGAGA GGCTGGACGG	1020
20	CTGTACAGCA TTTGCTCTGA TGTATGAGGG CCGCCGTGTG GCCATTCTTC GCAATCCAGA	1080
	GTTTTTTGAG CACAGGAAAG AGGAGCGCTG TGCCAGACAG TGGGGAACGA CATGCAAGAA	1140
25	CCACCCCTAT ATTAAGATGG TGATGGAACA AGGAGATTGG CTGATTGGAG GAGATCTTCA	1200
23	AGTCTTGGAT CGAGTTTATT GGAATGATGG TCTTGATCAG TATCGTCTTA CTCCTACTGA	1260
	GCTAAAGCAG AAATTTAAAG ATATGAATGC TGATGCTGTC TTTGCATTTC AACTACGCAA	1320
30	CCCAGTGCAC AATGGACATG CCCTGTTAAT GCAGGATACC CATAAGCAAC TTCTAGAGAG	1380
	GGGCTACCGG CGCCCTGTCC TCCTCCTCCA CCCTCTGGGT GGCTGGACAA AGGATGACGA	1440
35	TGTTCCTTTG ATGTGGCGTA TGAAGCAGCA TGCTGCAGTG TTGGAGGAAG GAGTTCTGAA	1500
55	TCCTGAGACG ACAGTGGTGG CCATCTTCCC ATCTCCCATG ATGTATGCTG GACCAACTGA	1560
	GGTCCAGTGG CATTGCAGAG CACGGATGGT TGCAGGAGCC AACTTTTACA TTGTTGGACG	1620
40	AGACCCTGCT GGCATGCCTC ATCCAGAAAC AGGGAAGGAT CTTTATGAGC CAAGTCATGG	1680
	TECCAAAGTG CTGACGATGG CCCCTGGTTT AATCACTTTG GAAATAGTTC CCTTTCGAGT	1740
45	TGCAGCTTAC AACAAGAAAA AGAAGCGTAT GGACTACTAT GACTCTGAAC ACCATGAAGA	1800
,,,	CTTTGAATTT ATTTCAGGAA CACGAATGCG CAAACTTGCT CGAGAAGGCC AGAAACCACC	1860
	TGAAGGTTTC ATGGCTCCCA AGGCTTGGAC CGTGCTGACA GAATACTACA AATCCTTGGA	1920
50	GAAAGCTTAG GCTGTTAACC CAGTCACTCC ACCTTTGACA CATTACTAGT AACAAGAGGG	1980
	GACCACATAG TCTCTGTGG CATTTCTTTG TGGTGTCTGT CTGGACATGC TTCCTAAAAA	2040
55	CAGACCATTT TCCTTAACTT GCATCAGTTT TGGTCTGCCT TATGAGTTCT GTTTTGAACA	2100
55	AGIGTAACAC ACTGATGGIT TTAATGTATC TTTTCCACTT ATTATAGTTA TATTCCTACA	2160
	ATACAATTTI AAAATIGICI TITTATATTA TATTTATGCT TCIGIGICAT GATTTITICA	2220
60	AGCTGTTATA TTAGTTGTAA CCAGTAGTAT TCACATTAAA TCTTGCTTTT TTTCCCCTTA	2280

	AAAAAAGAAA	AAAATTACCA	AACAATAAAC	TIGGCTAGAC	CTTGTTTTGA	GGATTTTACA	2340
. 5	AGACCTTTGT	AGCGATTAGA	TTTTTTTCT	ACATTGAAAA	TAGAAACTGC	TTCCTTTCTT	2400
,	CTTTCCAGTC	AGCTATTGGT	CTTTCCAGCT	GITATAATCT	AAAGTATTCT	TATGATCTGT	2460
	GTAAGCTCTG	AATGAACTTC	TTTACTCAAT	AAAATTAATT	TTTTGGCTTC	TTAAAAAAAT .	2520
10	Алалал	1		'	,		252 7

15 (2) INFORMATION FOR SEQ ID NO: 118:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1098 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 118:

_							
25	CGCATCACAG	ACAACCCAGA	AGGAAAATGG	TTGGGCAGAA	CAGCAAGGGG	TTCATATGGC	. 60
	AAAATTATAT	CAACTGCTGT	AGAGATTNNC	TATGATTCTT	TGAAACTGAA	AAAAGACTCT	120
30	CTTGGTGCCC	CTTCAAGACC	TATTGAAGAT	GACCAAGAAG	TATATGATGA	TGTTGCAGAG	180
	CAGGATGATA	TTAGCAGCCA	CAGTCAGAGT	GGAAGTGGAG	GGATATTCCC	TCCACCACCA	240
	GATGATGACA	TTTATGATGG	GATTGAAGAG	GAAGATGCTG	atgatggtt	CCCTGCTCCT	300
35	CCTAAACAAT	TGGACATGGG	AGATGAAGTT	TACGATGATG	TGGATACCTC	TGATTTCCCT	360
	GTTTCATCAG	CAGAGATGAG	TCAAGGAACT	AATGTTGGAA	AAGCTAAGAC	AGAAGAAAAG	420
40	GACCTTAAGA	AGCTAAAAA	GCAGRAAAA	GAARAAAAAG	ACTTCAGGAA	AAATTTAAA	480
40	TATGATGGTG	AAATTAGAGT	CCTATATTCA	ACTAAAGTTA	CAACTTCCAT	AACTTCTAAA	540
	AAGTGGGGAA	CCAGAGATCT	ACAGGTAAAA	CCTGGTGAAT	CTCTAGAAGT	TATACAAACC	600
45	ACAGATGACA	CAAAAGTTCT	CTGCAGAAAT	GAAGAAGGGA	AATATGGTTA	TGTCCTTCGG	660
	AGTTACCTAG	CGGACAATGA	TGGAGAGATC	TATGATGATA	TTGCTGATGG	CTGCATCTAT	720
50	GACAATGACT	AGCACTCAAC	TTTGGTCATT	CTGCTGTGTT	CATTAGGTGC	CAATGTGAAG	780
30	TCTGGATTTT	AATTGGCATG	TTATTGGGTA	TCMAGAAAAT	TAATGCACAR	AACCACTTAT	840
	TATCATTIGT	TATGAAATCC	CAATTATCTT	TACAAAGTGT	TTAAAGTTTG	AACATAGAAA	900
55	ATAATCTCTC	TGCTTAATTG	TTATCTCAGA	AGACTACATT	AGTGAGATGT	AAGAATTATT	960
		TTTCCGCTTT					1020
60	- WILLIAM	TAATCCTCCT	ICAAAAAAI'A	AAAATAAAAA	ААААААААА	AAACTCGAGG	1080

PCT/US98/05311 WO 98/42738

268

1098 GGGGCCCGG TACCCAAT

5

10

(2) INFORMATION FOR SEO ID NO: 119:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1679 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 119:

15 TOGACCCACG CGTCCGGCGA GATCCCTACC GCAGTAGCCG CCTCTGCCGC CGCGGAGCTT . 60 CCCGAACCTC TTCAGCCGCC CGGAGCCGCT CCCGGAGCCC GGCCGTAGAG GCTGCAATCG 120 CAGCCGGGAG CCCGCAGCCC GCGCCCCGAG CCCGCCGCCG CCCTTCGAGG GCGCCCCAGG 20 180 CCGCGCCATG GTGAAGGTGA CGTTCAACTC CGCTCTGGCC CAGAAGGAGG CCAAGAAGGA 240 CGAGCCCAAG AGCGGCGAGG AGGCGCTCAT CATCCCCCCC GACGCCGTCG CGGTGGACTG 300 25 CAAGGACCCA GATGATGTGG TACCAGTTGG CCAAAGAAGA GCCTGGTGTT GGTGCATGTG 360 CTTTGGACTA GCATTTATGC TTGCAGGTGT TATTCTAGGA GGAGCATACT TGTACAAATA 420 30 TTTTGCACTT CAACCAGATG ACGTGTACTA CTGTGGAATA AAGTACATCA AAGATGATGT 480 CATCTTAAAT GAGCCCTCTG CAGATGCCCC AGCTGCTCTC TACCAGACAA TTGAAGAAAA 540 TATTAAAATC TTTGAAGAAG AAGAAGTTGA ATTTATCAGT GTGCCTGTCC CAGAGTTTGC 600 35 660 AGATAGTGAT CCTGCCAACA TTGTTCATGA CTTTAACAAG AAACTTACAG CCTATTTAGA TCTTAACCTG GATAAGTGCT ATGTGATCCC TCTGAACACT TCCATTGTTA TGCCACCCAG 720 40 AAACCTACTG GAGTTACTTA TTAACATCAA GGCTGGAACC TATTTGCCTC AGTCCTATCT 780 GATTCATGAG CACATGGTTA TTACTGATCG CATTGAAAAC ATTGATCACC TGGGTTTCTT 840 900 TATTTATCGA CTGTGTCATG ACAAGGAAAC TTACAAACTG CAACGCAGAG AAACTATTAA 45 960 AGGTATTCAG AAACGTGAAG CCAGCAATTG TTTCGCAATT CGGCATTTTG AAAACAAATT 1020 TGCCGTGGAA ACTITAATTT GTTCTTGAAC AGTCAAGAAA AACATTATTG AGGAAAATTA 50 1080 ATATCACAGC ATAACCCCAC CCTTTACATT TTGTGCAGTG ATTATTTTTT AAAGTCTTCT 1140 TTCATGTAAG TAGCAAACAG GGCTTTACTA TCTTTTCATC TCATTAATTC AATTAAAACC ATTACCTTAA AATTITTITC TITCGAAGTG TGGTGTCTTT TATATTTGAA TTAGTAACTG 1200 55 1260 TATGAAGTCA TAGATAATAG TACATGTCAC CTTAGGTAGT AGGAAGAATT ACAATTTCTT 1320 TAAATCATTT ATCTGGATTT TTATGTTTTA TTAGCATTTT CAAGAAGACG GATTATCTAG 60 1380 AGAATAATCA TATATATGCA TACGTAAAAA TGGACCACAG TGACTTATTT GTAGTTGTTA

	GTTGCCCTGC	TACCTAGITT	GTTAGTGCAT	TTGAGCACAC	ATTTTAATTT.	TCCTCTAATT	1440
_	AAAATGTGCA	GTATTTTCAG	TGTCAAATAT	ATTTAACTAT	TTAGAGAATG	ATTTCCACCT	1500
5	TTATGTTTTA	ATATCCTAGG	CATCTGCTGT	AATAATATTT	TAGAAAATGT	TTGGAATTTA	1560
	AGAAATAACT	TGTGTTACTA	ATTTGTATAA	CCCATATCTG	TGCAATGGAA	TATAAATATC	1620
10	ACAAAGITGT	TTAAMWAAAA	ААААААААА	ААААААААА	ААААААААА	AAAAAAA	1679

15 (2) INFORMATION FOR SEQ ID NO: 120:

20

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1308 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear '

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 120:

25	TTGGCANCNG GGAGAGGGAA AGAGGAGGAA ATGGGGTTTG AGGACCATGG CTTACCTTTC	60
	CTGCCTTTGA CCCATCACAC CCCATTTCCT CCTCTTTCCC TCTCCCCGCT GCCAAAAAAA	120
30	AAAAAAAAGG AAACGTTTAT CATGAATCAA CAGGGTTTCA GTCCTTATCA AAGAGAGATG	180
30	TGGAAAGAGC TAAAGAAACC ACCCTTTGTT CCCAACTCCA CTTTACCCAT ATTTTATGCA	240
	ACACAAACAC TGTCCTTTTG GGTCCCTTTC TTACAGATGG ACCTCTTGAG AAGAATTATC	300
35	GTATTCCACG TTTTTAGCCC TCAGGTTACC AAGATAAATA TATGTATATA TAACCTTTAT	360
	TATTGCTATA TCTTTGTGGA TAATACATTC AGGTGGTGCT GGGTGATTTA TTATAATCTG	420
40	AACCTAGGTA TATCCTTTGG TCTTCCACAG TCATGTTGAG GTGGGCTCCC TGGTATGGTA	480
	AAAAGCCAGG TATAATGTAA CTTCACCCCA GCCTTTGTAC TAAGCTCTTG ATAGTGGATA	540
	TACTCTTTTA AGTTTAGCCC CAATATAGGG TAATGGAAAT TTCCTGCCCT CTGGGTTCCC	600
45	CATTITIACT ATTAAGAAGA CCAGIGATAA TITAATAATG CCACCAACTC TGGCTTAGIT	660
	AAGTGAGAGT GTGAACTGTG TGGCAAGAGA GCCTCACACC TCACTAGGTG CAGAGAGCCC	720
50	AGGCCTTATG TTAAAATCAT GCACTTGAAA AGCAAACCTT AATCTGCAAA GACAGCAGCA	780
50	AGCATTATAC GGTCATCTTG AATGATCCCT TTGAAATTTT TTTTTTGTTT GTTTGTTTAA	840
	ATCAAGCCTG AGGCTGGTGA ACAGTAGCTA CACACCCATA TTGTGTGTTC TGTGAATGCT	900
55	AGCTCTCTTG AATTTGGATA TTGGTTATTT TTTATAGAGT GTAAACCAAG TTTTATATTC	960
	TGCAATGCGA ACAGGTACCT ATCTGTTTCT AAATAAAACT GTTTACATTC ATTATGGGGT	1020
60	ATGTATGACC TTCATTTTCC AAGAAATAGA ACTCTAGCTT AGAATTATGG ATGCTCTAAA	1080

ATGTCAGAAT GGGAACTCTC CTCGAAGTTC TCCCAAACTC AGAGACAGCA CTGCCTTCTC 1140 CTAAATGATT ATTCTTTCT CCCTGTTTTC TGGTATTTTC TAGGCATCCT TCTCACCACA 1200 5 GCCATAACCC TTTTTTACTT CCATTAGGCC GTATAACTGG NGGGACNGCT GGTCGGTATA 1260 TAATACTGGT WCCAACAMAG GGGTTCTGGA TGTACACMAG GTTATCTT 1308 (2) INFORMATION FOR SEQ ID NO: 121:

10

15

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1411 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 121:

GGCACAGGAG CGACCCGGGA GAAGGAGGGC CAMGAKGCGG AAGCGGAGGA GTCTCCAGGA 60 GACCCGGGGA CAGCATCGCC CAGGCCCCTG TTTGCAGGCC TTTCAGATAT ATCCATCTCA 120 25 CAAGACATCC CCGTAGAAGG AGAAATCACC ATTCCTATGA GATCTCGCAT CCGGGAGTTT 180 GACAGCTCCA CATTAAATGA ATCTGTTCGC AATACCATCA TGCGTGATCT AAAAGCTGTT 30 GGGAAAAAAT TCATGCATGT TTTGTACCCA AGGAAAAGTA ATACTCTTTT GAGAGATTGG 300 GATTTGTGGG GCCCTTTGAT CCTTTGTGTG ACACTCGCAT TAATGCTGCA AAGAGACTCT 360 GCAGATAGTG AAAAAGATGG AGGGCCCCAA TTTGCAGAGG TGTTTGTCAT TGTCTGGTTT 420 35 GGTGCAGTTA CCATCACCCT CAACTCAAAA CTTCTTGGAG GGAACATATC TFFFFTTTCAG 480 AGCCTCTGTG TGCTGGGTTA CTGTATACTT CCCTTGACAG TAGCAATGCT GATTTGCCGG 540 40 CTGGTACTTT TGGCTGATCC AGGACCTGTA AACTTCATGG TTCGGCTTTT TGTGGTGATT 600 GTGATGTTTG CCTGGTCTAT AGTTGCCTCC ACAGCTTTCC TTGCTGATAG CCAGCCTCCA 660 AACCGCAGAG CCCTAGCTGT TTATCCTGTT TTCCTGTTTT ACTTTGTCAT CAGTTGGATG 720 45 ATTCTCACCT TTACTCCTCA GTAAATCAGG AATGGGAAAT TAAAAACCAG TGAATTGAAA 780 GCACATCTGA AAGATGCAAT TCACCATGGA GCTTTGTCTC TGGCCCTTAT TTGTCTAATT 840 50 TTGGAGGTAT TTGATAACTG AGTAGGTGAG GAGATTAAAA GGGAGCCATA TAGCACTGTC 900 ACCCCITATT TGAGGAACTG ATGTTTGAAA GGCTGTTCTT TTCTCTCTTA ATGTCATTTC 960 TTTAAAAATA CATGTGCATA CTACACACAG TATATAATGC CTCCTTAAGG CATGATGGAG 1020 55 TCACCGTGGT CCATTTGGGT GACAACCAGT GACTTGGGAA GCACATAGAT ACATCTTACA 1080 AGTTGAATAG AGTTGATAAC TATTTTCAGT TTTGAGAATA CCAGTTCAGG TGCAGCTCTT 1140 60 AAACACATTG CCTTATGACT ATTAGAATAT GCCTCTCTTT TCATAAATAA AAATACATGG 1200

	TCTATATCCA THITCTITTA TITCTCTCTC TTAAGCTTAA AAAGGCAATG AGAGAGGTTA	1260
5	GGAGTGGGTT CATACACGGA GAATGAGAAA ACATGCATTA ACCAATATTC AGATTTTGAT	1320
3	CAGGGGAAAT TCTAYACTTG TTGCAAAAAA AAAAAAAAA AAACTCGAGG GGGCCCCGT	1380
	ACCCAATCGC NGTATATGAT CGNAAACAAT C	1411
10		
	(2) INFORMATION FOR SEQ ID NO: 122:	
15 20	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 2256 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	•
20	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 122:	
	GCTTTGGCTT TTTTTGGCGG ACTGGGGCGC CCTCCGGAAG CGTTTCCAAC TTTCCAGAAG	60
25	TTTCTCGGGA CGGGCAGGAG GGGGTGGGGA CTGCCATATA TAGATCCCGG GAGCAGGGGA	120
	GCGGGCTAAG AGTAGAATCG TGTCGCGGCT CGAGAGCGAG AGTCACGTCC CGGCGCTAGC	180
30	CAGCCCGACC CAGGCCCACC GTGGTGCACG CAAACCACTT CCTGGCCATG CGCTCCCTCC	240
30	TECTTCTCAG CGCCTTCTGC CTCCTGGAGG CGGCCCTGGC CGCCGAGGTG AAGAAACCTG	300
	CAGCCGCAGC AGCTCCTGGC ACTGCGGAGA AGTTGAGCCC CAAGGCGGCC ACGCTTGCCG	360
35	AGCGCANGCC GGCCTGGCCT TCAGCTTGTA CCAGGCCATG GCCAAGGACC AGGCAGTGGA	420
	GAACATCCTG GTGTCACCCG TGGTGGTGGC CTCGTCGCTG GGGCTCGTGT CGCTGGGCGG	480
40	CAAGGCGACC ACGGCGTCGC AGGCCAAGGC AGTGCTGAGC GCCGAGCAGC TGCGCGACGA	540
40	GGAGGTGCAC GCCGGCCTGG GCGAGCTGCT GCGCTCACTC AGCAACTCGA CGGCGCGCAA	600
	CGTGACCTGG AAGCTGGGCA GCCGACTGTA CGGACCCAGC TCAGTGAGCT TCGCTGATGA	660
45	CTTCGTGCGC ACAGCAAGCA GCACTACAAC TGCGAGCACT CCAAGATCAA CTTCCGCGAC	720
	AAGCGCAGNG CGCTGCAGTC CATCAACGAG TGGGCCGCGC AGACCACCGA CGGCAAGCTG	780
50	CCCGAGGTCA CCAAGGACGT GGAGCGCACG GACGGCGCCC TGCTAGTCAA CGCCATGTTC	840
30	TTCAAGCCAC ACTGGGATGA GAAATTCCAC CACAAGATGG TGGACAACCG TGGCTTCATG	900
	GTGACTCGGT CCTATACYGT GGGTGTCATG ATGATGCACC GGACAGGCCT CTACAACTAC	960
55	TACGACGACG AGAAGGAAAA GCTGCAAATC GTGGAGATGC CCCTGGCCCA CAAGCTCTCC	102
	AGCCTCATCA TCCTCATGCC CCATCACGTG GAGCCTCTCG AGCGCCTTGA AAAGCTGCTA	108
60	ACCAAAGAGC AGCTGAAGAT CTGGATGGGG AAGATGCAGA AGAAGGCTGT TGCCATCTCC	114

	TTGCCCAAGG GTGTGGTGGA GGTGACCCAT GACCTGCAGA AACACCTGGC TGGGCTGGGC	1200
	CTGACTGAGG CCATTGACAA GAACAAGGCC GACTTRTCAC GCATGTCAGG CAAGAAGGAC	1260
5	CTGTACCTGG CCAGCGTGTT CCACGCCACC GCCTTTGAGT TGGACACAGA TGGCAACCCC	1320
	TTTGACCAGG ACATCTACGG GCGCGAGGAG CTGCGCANCC CAAGCTGTTC TACGCCGACC	1380
10	ACCCCITCAT CTTCCTAGIG CGGGACACCC AAAGCGGCTC CCTGCTATIC ATTGGGCGCC	1440
	TGGTCCGGCC TAAGGGTGAC AAGATGCGAG ACGAGTTATA GGGCCTCAGG GTGCACACAG	1500
	GATGGCAGGA GGCATCCAAA GGCTCCTGAG ACACATGGGT GCTATTGGGG TTGGGGGGGA	1560
15	GGTGAGGTAC CAGCCTTGGA TACTCCATGG GGTGGGGGTG GAAAARCAGA CCGGGGTTCC	1620
	CGTGTGCCTG AGCGGACCTT CCCAGCTAGA ATTCACTCCA CTTGGACATG GGCCCCAGAT	1680
20	ACCATGATGC TGAGCCCGGA AACTCCACAT CCTGTGGGAC CTGGGCCATA GTCATTCTGC	1740
	CTGCCCTGAA AGTCCCAGAT CAAGCCTGCC TCAATCAGTA TTCATATTA TAGCCAGGTA	1800
	CCTTCTCACC TGTGAGACCA AATTGAGCTA GGGGGGTCAG CCAGCCCTCT TCTGACACTA	1860
25	AAACACCTCA GCTGCCTCCC CAGCTCTATC CCAACCTCTC CCAACTATAA AACTAGGTGC	1920
	TGCAGCCCCT GGGACCAGGC ACCCCCAGAA TGACCTGGCC GCAGTGAGGC GGATTGAGAA	1980
30	GGAGCTCCCA GGAGGGGCTT CTGGGCAGAC TCTGGTCAAG AAGCATCGTG TCTGGCGTTG	2040
	TGGGGATGAA CTTTTTGTTT TGTTTCTTCC TTTTTTAGTT CTTCAAAGAT AGGGAGGGAA	2100
	GGGGGAACAT GAGCCTTTGT TGCTATCAAT CCAAGAACTT ATTTGTACAT TTTTTTTTTC	2160
35	AATAAAACTT TTCCAATGAC AAAAAAAAAA AAAAAAAAA AAAAAGGGGS GGGCCGCTCC	2220
	TAGAGGGATC CCTCCGANGG NGCCCAATCG AAAATN	2256
40		
	(2) INFORMATION FOR SEQ ID NO: 123:	
		•
45	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 829 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double (D) TOPOLOGY: linear	
50	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 123:	
	ATGCGCTCCC TCCTGCTTCT CAGCGCCTTC TGCCTCCTGG AGGCGGCCCT GGCCGCCGAG	60
		60
55	GTGAAGAAAC CTGCAGCCGC AGCAGCTCCT GGCACTGCGG AGAAGTTGAG CCCCAAGGCG	120
	GCCACGCTTG CCGAGCGEAA GCGGCCTGGC CTTCAGCTTG TACCAGGCCA TGGCCAAGGA	180
	CCAGGCAGIG GAGAACATCC TGGTGTCACC CGTGGTGGTG GCCTCGTCGC TGGGGCTCGT	240
60	GTCGCTGGGC GGCAAGGCGA CCACGGCGTC GCAGGCCAAG GCAGTGCTGA GCGCCGAGCA	300

	GCTGCGCGAC GAGGAGGTGC ACGCCGGCCT GGGCGAGCTG CTGCGCTCAC TCAGCAACTC	360
5	CACGGCGCG AACGTGACCT GGAAGCTGGG CAGCCGACTG TACGGACCCA GCTCAGTGAG	420
3	CTTCGCTGAT GACTTCGTGC GCAGCAGCAA GCAGCACTAC AACTGCGAGC ACTCCAAGAT	480
	CAACTTCCGC GACAAGCGCA GCGCGCTGCA GTCCATCAAC GAGTGGGCCG CGCAGACCAC	540
10	CGACGGCAAG CTGCCGGAGG TCACCAAGGA CGTGGAGGGC ACGGACGGCG CCCTGTTAGT	600
	CAACGCCATG TTCTTCAAGC CACACTGGGA TGAGAAATTC CACCACAAGA TGGTGGACAA	660
15	CCGTGGCTTC ATGGTGACTC GGTCCTATAC CGTGGGTGTC ATGATGATGC ACCGGACAGG	720
15	CCTCTACAAC TACTACGACG ACGAGAAGGA AAAGCTGCAA ATCGTGGAGA TGCCCCTGGC	780
	CCACAAGCTC TCCAGCCTCA TCATCCTCAT GCCCCATCAC GTGGAGCCT	829
20		
	(2) INFORMATION FOR SEQ ID NO: 124:	
	(2) INCOMPLETION FOR SING 25 NO. 222.	
25	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 2223 base pairs	-
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double (D) TOPOLOGY: linear	
30	(5) 101030011 123002	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 124:	
	CCTCCGGAAG CGTTTCCAAC TTTCCAGAAG TTTCTCGGGA CGGCAGGAG GGGGTGGGGA	60
35	CTGCCATATA TAGATCCCGG GAGCAGGGGA GCGGGCTAAG AGTAGAATCG TGTCGCGGCT	120
	CGAGAGCGAG AGTCACGTCC CGGCGCTAGC CAGGCCCCACC GTGGTGCACG	180
40	CAAACCACTT CCTGGCCATG CGCTCCCTCC TGCTTCTCAG CGCCTTCTGC CTCCTGGAGG	240
.0	CGGCCCTGGC CGCCGAGGTG AAGAAACCTG CAGCCGCAGC AGCTCCTGGC ACTGCGGAGA	300
	AGTTGAGCCC CAAGGCGGCC ACGCTTGCCG AGCGCAGNCG GCCTGGCCTT CAGCTTGTAC	360
45	CAGGCCATGG CCAAGGACCA GGCAGTGGAG AACATCCTGG TGTCACCCGT GGTGGTGGCC	420
	TOGTOGOTGG GGCTCGTGTC GCTGGGCGGC AAGGCGACCA CGGCGTCGCA GGCCAAGGCA	480
50	GTGCTGAGCG CCGAGCAGCT GCGCGACGAG GAGGTGCACG CCGGCCTGGG CGAGCTGCTG	540
	CGCTCACTCA GCAACTCSAC GGCGCGCAAC GTGACCTGGA AGCTGGGCAG CCGACTGTAC	600
	GGACCCAGCT CAGTGAGCTT CGCTGATGAC TTCGTGCGCA CAGCAAGCAG CACTACAACT	660
55	GCGAGCACTC CAAGATCAAC TTCCGCGACA AGCGCACGCG CTGCAGTCCA TCAACGAGTG	720
	GGCCGCCCAG ACCACCGACG GCAAGCTGCC CGAGGTCACC AAGGACGTGG AGCGCACGGA	780
	COCCOCCOMO ANTACTICA ACO CONTINUENTO CARCOCACAC TOCCOMORACA ARTITICACIA	940

	CAAGATGGTG	GACAACCGTG	GCTTCATGGT	GACTCGGTCC	TATACYGTGG	GIGTCATGAT	900
	GATGCACCGG	ACAGGCCTCT	ACAACTACTA	CGACGACGAG	AAGGAAAAGC	TGCAAATCGT	960
5	GGAGATGCCC	CTGGCCCACA	AGCTCTCCAG	CCTCATCATC	CTCATGCCCC	ATCACGTGGA	1020
	GCCTCTCGAG	CGCCTTGAAA	AGCTGCTAAC	CAAAGAGCAG	CTGAAGATCT	GGATGGGGAA	1080
10	GATGCAGAAG	AAGGCTGTTG	CCATCTCCTT	GCCCAAGGGT	CTCCTCGACG	TGACCCATGA	1140
	CCTGCAGAAA	CACCIGGCIG	GGCTGGGCCT	GACTGAGGCC	ATTGACAAGA	ACAAGGCCGA	1200
	CTTRTCACGC	ATGTCAGGCA	AGAAGGACCT	GTACCTGGCC	ACCOTOTTCC	ACGCCACCGC	1260
15	CTTTGAGTTG	GACACAGATG	GCAACCCCTT	TGACCAGGAC	ATCTACGGGC	GCGAGGAGCT	1320
	GCGCASCCCA	AGCTGTTCTA	CGCCGACCAC	CCCTTCATCT	TCCTACTGCG	GGACACCCAA	1380
20	AGCGGCTCCC	TGCTATTCAT	TEGECECCTE	GTCCGGCCTA	AGGGTGACAA	GATGCGAGAC	1440
	GAGTTATAGG	GCCTCAGGGT	GCACACAGGA	TGGCAGGAGG	CATCCAAAGG	CTCCTGAGAC	1500
	ACATGGGTGC	TATTGGGGTT	GGGGGGAGG	TGAGGTACCA	GCCTTGGATA	CTCCATGGGG	1560
25	TGGGGGTGGA	AAARCAGACC	GGGTTCCCG	TGTGCCTGAG	CGGACCTTCC	CAGCTAGAAT	1620
	TCACTCCACT	TGGACATGGG	CCCCAGATAC	CATGATGCTG	AGCCCGGAAA	CTCCACATCC	1680
30	TGTGGGACCT	GGGCCATAGT	CATTCTGCCT	GCCCTGAAAG	TCCCAGATCA	AGCCTGCCTC	1740
	AATCAGTATT	CATATTTATA	GCCAGGTACC	TTCTCACCTG	TGAGACCAAA	TTGAGCTAGG	1800
	GGGGTCAGCC	AGCCCTCTTC	TGACACTAAA	ACACCTCAGC	TGCCTCCCCA	GCTCTATCCC	1860
35	AACCTCTCCC	AACTATAAAA	CTAGGTGCTG	CAGCCCCTGG	GACCAGGCAC	CCCCAGAATG	1920
	ACCTGGCCGC	AGTGAGGCGG	ATTGAGAAGG	AGCTCCCAGG	AGGGGCTTCT	GGGCAGACTC	1980
10	TGGTCAAGAA	GCATCGTGTC	TOGCCTTCTC	GGGATGAACT	TTTTGTTTTG	TTTCTTCCTT	2040
	TTTTAGTTCT	TCAAAGATAG	GGAGGGAAGG	GGGAACATGA	CCTTTCTTC	CTATCAATCC	2100
	AAGAACTTAT	TTGTACATTT	TTTTTTCAA	TAAAACTTTT	CCAATGACAA	AAAAAAAA	2160
15	Алалалала	MWMGGGGSGG	GCCGCTCCTA	GAGGGATCCC	TCCGANGGNG	CCCAATCGAA	2220
	AAT						2223

55

(2) INFORMATION FOR SEQ ID NO: 125:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 125:

60 Met Lys Lys Gln Ser Lys Arg Cys Leu Trp Lys Pro Pro Gly Ser Leu

	1			•	5				,	10					15	
5	Arg	Arg	Leu	Trp 20	Trp	Met	Arg	Ala	Leu 25	Leu	Ile	Leu	Lys	Туг 30	Ile	
	(2)		22242	n=01	500	orno.	·	70 1		•			ŧ			
	(2)	TNF	ORMAI	LION	FOR	SEQ	TD I	W: 1	126:							
10			(i) :	() ()	A) L B) T D) T	engt YPE: OPOL	H: 4 ami OGY:	5 am no a lin	ino cid ear	acid		! 12	' 6:	٠		
15	Met 1	Lys	Lys	Ser	Leu 5	Glu	Asn	Leu	Asn	Arg 10	Leu	Gln	Val	Met	Leu 15	Leu
20	His	Leu	Thr	Ala 20	Ala	Phe	Leu	Gln	Arg 25	Ala	His	Xaa	Ile	Leu 30	Thr	Thr
	Arg	Met	Ser 35	Leu	Gly	Phe	Gln	Ser 40	Pro	His	Leu	Thr	Met 45			
25	(2)	INF	ORMA!	PION	FOR	SEQ	ID I	vo: :	127:							
			(i)	SEOU	ENCE	СНА	RACT	ERIS	TTCS							
30				((A) L B) T D) T	ENGT YPE : OPOL	H: 3 ami OGY:	9 am no a lin PTIO	ino .cid .ear	acid		: 12	7:			ı
35	Met 1		Asn	Gln	Arg 5	Gln	Val	Phe	Leu	Phe 10	His	Leu	Phe	Ser	Asn 15	Тут
40	Leu	Leu	Ser	Ile 20	Asn	Ser	Val	Pro	Gly 25		Leu	Leu	Ala	Ala 30	Thr	Tyr
	Cys	Leu	Asn 35	Met	Thr	Tyr	Gly									
45	(2)	INF	ORMA	TION	FOR	SEQ	ID:	NO:	128:							
			(i)					ERIS 3 an			ls					
50			(xi)	(B) I	YPE:	ami OGY :	no a lir PTIC	icid near): 12	:8:			
55	Met 1	_	Lys	Lys	Phe 5		Leu	Ala	Gln	Val		Leu	Ser	Leu	Ser 15	
	Met	Pro	Ser	Met 20		Val	Thr									
60																

	(2)	INF	ORMA	TION	FOR	SEQ	ID	NO:	129:							
. 2				(ENCE (A) I (B) I (D) I	ENGI YPE : OPOL	H: 1 ami OGY:	no a	mind cid ear	aci	.ds	, ' : 12	9:			
10	Met 1		Leu	Leu	Cys 5	Leu	Leu	Leu	Val	Pro 10	Ĺeu	Leu	Leu	Ser	Leu 15	Phe
15	Val	Leu	Gly	Leu 20	Phe	Leu '	Trp	Phe	Leu 25	Lys	Arg	Glu	Arg	Gln 30	Glu	Glu
	Tyr	Ile	Glu 35	Glu	Lys	Lys	Arg	Val 40	Asp	Ile	Çys	Arg	Glu 45	Thr	Pro	Asr
20	Ile	Суs 50	Pro	His	Ser	Gly ,	Glu 55	Asn	Thr	Glu	Tyr	Asp 60	Thr	Ile	Pro	His
	Thr 65	Asn	Arg	Thr	Ile	Leu 70	Lys	Glu	Asp	Pro	Ala 75	Asn	Thr	Val	Tyr '	Ser 80
25	Thr	Val	Glu	Ile	Pro 85	Lys	Lys	Met	Glu	Asn 90	Pro	His	Ser	Leu	Leu 95	Thr
30	Met	Pro	Asp	Thr 100	Pro	Arg	Leu	Phe	Ala 105	Tyr	Glu	Asn	Va1	Ile 110		
35	(2)	INF			FOR ENCE							1 1				
				(A) L B) T D) T UENCI	ENGT YPE : OPOL	H: 6 ami OGY:	3 am no a lin	ino d cid ear	acid		. 12	٥.			
40																
	Met 1	Leu	Leu	Leu	Phe 5	Ile	Tyr	Phe	Tyr	Ser 10	His	Pro	Ala	Pro	Val 15	Pro
45	Ala	Gly	Ala	Thr 20	Ser	Lys	Pro	Arg	Тут 25	Arg	Val	Ile	Thr	Суs 30	Gly	Pro
	Ala	Ser	Val 35	Phe	Ser	Thr	Ser	Phe 40	Ser	His	Ser	Pro	Pro 45	Ala	Arg	Cys
50	Leu	Gly 50	Arg	Leu	Glu	Gln	Met 55	Phe	His	Phe	Gly į	Leu 60	Ala	Ser	Gly	
55	(2)	INFO	ORMAI	MOI	FOR	SEQ	ID N	ю: 1	.31:							
			(i) :	(;	ENCE A) LI	NGT	H: 3	0 am	ino a		5					
60					B) TY D) TY											

```
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 131:
     Met Pro Phe Pro Ile Ser Ile Leu Gln Leu Cys Leu Gln Ile Ser Asn
                                          10
5
     Leu Ser Phe Cys Leu Gln Lys Ile Tyr Lys Ile Pro Phe Val
                                      25 '
10
     (2) INFORMATION FOR SEQ ID NO: 132:
             (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 53 amino acids
15
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 132:
     Met Ala Ala Cys Arg Ser Val Lys Gly Leu Val Ala Val Ile Thr
20
                                           10
     Gly Gly Ala Ser Gly Leu Gly Leu Ala Thr Ala Asp Asp Leu Trp Gly
      Arg Glu Pro Leu Leu Cys Phe, Trp Thr Cys Pro Thr Arg Val Gly Arg
25
               35
      Pro Lys Pro Arg Ser
           50
30
      (2) INFORMATION FOR SEQ ID NO: 133:
              (i) SEQUENCE CHARACTERISTICS:
35
                     (A) LENGTH: 57 amino acids
                     (B) TYPE: amino acid
                     (D) TOPOLOGY: linear
              (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 133:
40
      Met Leu Leu Val Tyr Asp Leu Tyr Leu Xaa Pro Lys Leu Trp Ala Leu
                                            10
      Ala Thr Pro Gln Lys Asn Gly Lys Gly Ala Arg Xaa Gly Asp Gly Thr
 45
                                        25
       Pro Ala Gln Ala Phe Trp Asp Phe Trp Ser His Leu Ile Ser Ala Asp
 50
       Pro Gln Thr Trp Glu Arg Ala Ala Pro
 55
       (2) INFORMATION FOR SEQ ID NO: 134:
              (i) SEQUENCE CHARACTERISTICS:
                      (A) LENGTH: 216 amino acids
                      (B) TYPE: amino acid
                      (D) TOPOLOGY: linear
 60
```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 134:

· 5	Met 1	Arg	, Leu	Ser	Ala 5	Leu	Leu	Ala	Leu	Ala 10		Lys	Val	Thr	Leu 15	
	Pro	His	Tyr	Arg 20	Tyr	Gly	Met	Ser	Pro 25		Gly	Ser	Val	Ala 30		Lys
10	Arg	i Lys	Asn 35	Pro	Pro	Trp	'Ile	Arg 40	Arg	Arg	Pro	Val	Val 45		Glu	Pro
	Ile	Ser 50	Asp	Glu	Asp	Trp	Tyr 55	Leu	Phe	Cys	Gly	Asp 60		Val	Glu	Ile
15	Leu 65	Glu	Gly	Lys	Asp	Ala 70	Gly 	Lys	Gln	Gly	Lys 75	Val	Val	Gln	Val	Ile 80
20	Arg	Gln	Arg	Asn	Trp 85	Val	Val	Val	Gly	Gly 90	Leu	Asn	Thr	His	Tyr 95	Arg
	Tyr	Ile	Gly	Lys 100	Thr	Met	Asp	Tyr	Arg 105	Gly	Thr	Met	Ile	Pro 110	Ser	Glu
25	Ala	Pro	Leu 115	Leu	His	Arg	Gln	Val 120	Lys	Leu	Val	Asp	Pro 125	Met	Asp	Arg
	Lys	Pro 130	Thr	Glu	Ile	Glu	Trp 135	Arg	Phe	Thr	Glu	Ala 140	Gly	Glu	Arg	Val
30	Arg 145	Val	Ser	Thr	Arg	Ser 150	Gly	Arg	Ile	Ile	Pro 155	Lys	Pro	Glu	Phe	Pro 160
35	Arg	Ala	Asp	Gly	Ile 165	Val	Pro	Glu	Thr	Trp 170	Ile	Asp	Gly	Pro	Lys 175	Asp
	Thr	Ser	Val	Glu 180	Asp	Ala	Leu	Glu	Arg 185	Thr	Tyr	Val	Pro	Cys 190	Leu	Lys
40	Thr	Leu	Gln 195	Glu	Glu	Val	Met	Glu 200	Ala	Met	Gly	Ile	Lys 205	Glu	Thr	Arg
	Lys	Туг 210	Lys	Lys	Val	Tyr	Trp 215	Tyr								
45	(2)	INFO	RMAT	'ION	FOR	SEO	ID N	iO: 1	35.							
					NCE											
50				(<i>I</i>	A) LE B) TY D) TO	NGTI PE:	I: 49 amir	ami no ac	ino a		3					
		((xi)		ENCE					D II	NO:	135	i :			
55	Met 1	Ser	Leu	Arg	Gln i 5	Lys	Ser	Ser	Phe	Arg 10	Leu	Met	Val	Met	Ser 15	Leu
	Thr	Ile	Leu		Leu :	Ser	Lys	Thr		Val :	Leu	Cys	Leu		Cys	Leu
60				20					25					30		

```
His Ser Leu Lys Leu Thr Trp Arg Asp Gly Ala Arg Cys Ile Asn Ala
     Glu
 5
      (2) INFORMATION FOR SEQ ID NO: 136:
10
                     1
             (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 68 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
15
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 136:
      Met Ser Gly Ser Phe Ile Leu Cys Leu Ala Leu Val Thr Arg Trp Ser
20
      Pro Gln Ala Ser Ser Val 'Pro Leu Ala Val Tyr Glu Ser Lys Thr Arg
                                       25 '
      Lys Ser Tyr Arg Ser Gln Arg Asp Arg Asp Gly Lys Asp Arg Ser Gln
                                   40
25
      Gly Met Gly Leu Ser Leu Leu Val Glu Thr Arg Lys Leu Leu Leu Ser
                                                   60
      Ala Asn Gln Gly
30
       65
      (2) INFORMATION FOR SEQ ID NO: 137:
35
             (i) SEQUENCE CHARACTERISTICS:
                     (A) LENGTH: 52 amino acids
                     (B) TYPE: amino acid
                     (D) TOPOLOGY: linear
40
              (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 137:
      Met Cys Phe Arg Phe Phe Leu Phe Cys Ser Arg Ile Leu Leu Lys Leu
        1
                        5
45
      Phe Phe Leu Leu Phe Pro Ala Ser Ala Phe Pro Leu Ser Thr Arg Ser
      Ser Leu Ser Val Asn Glu His Val Val Val Ser Pro Arg Ser Thr Val
                                                        45
50
      Ser Ile Ser Arg
           50
55
       (2) INFORMATION FOR SEQ ID NO: 138:
              (i) SEQUENCE CHARACTERISTICS:
                     (A) LENGTH: 541 amino acids
60
                     (B) TYPE: amino acid
```

(D) TOPOLOGY: linear, (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 138:

5	Met	t Va l	l Ar	g Th	r Ası	o G1;	y Hi:	5 Thi	: Le	u Se:		ı Ly	s Ar	g Ası	n Ty	r Glr 5
	Val	l Th	r As	n Se 2	r Met 0	: Pho	e Gly	/ Ala	Sei 25	r Arg	g Lys	S Ly:	s Phe	e Val 30		u Gly
10	Va]	L As _l	Se:	r As _l 5	р Туг	His	s Ası	Glu 40	Ası	n Met	Туг	Ту	Sei 45	_	ı Se	r Ser
15	Met	: Phe 50	e Pro	O His	s Arg	Sez	Glu 55	Lys	Asp) Met	: Lev	Ala 60		Pro	Sei	Thr
	Ser 65	Gly	/ Gli	ı Let	ı Ser	Glr 70	n Phe	: Gly	Ala	Ser	: Leu 75		Gly	Gln	Glr	ser 80
20	Ala	Lev	ı Gly	/ Let	Pro 85	Met	Arg	Gly	Met	Ser 90		Asr	Thr	Pro	Glr 95	Leu
	Asn	Arg	Ser	100	Ser	Gln	Gly	Thr	Gln 105		Pro	Ser	His	Val 110		Pro
25	Thr	Thr	Gly 115	Val	. Pro	Thr	Met	Ser 120	Leu	His	Thr	Pro	Pro 125	Ser	Pro	Ser
30	Arg	Gly 130	Ile	: Leu	Pro	Met	Asn 135	Pro	Xaa	Asn	Met	Met 140		His	Ser	Gln
	143				Ile	150					155					160
35					Ser 165					170					175	
40				180	Pro				185					190		
40			133		Asn			200					205			
45		210			Phe		215					220				
	Leu 225	Asp	Leu	Ser	Asp	Phe 230	Pro	Ala	Leu	Ala	Asp 235	Arg	Asn	Arg	Arg	Glu 240
50	Gly	Ser	Gly	Asn	Pro 245	Thr	Pro	Leu	Ile	Asn 250	Pro	Leu	Ala	Gly	Arg 255	Ala
	Pro	Tyr	Val	Gly 260	Met '	Val	Thr		Pro 265	Ala	Asn	Glu		Ser 270	Gln	Asp
55	Phe	Ser	Ile 275	His	Asn (Glu	Asp	Phe : 280	Pro	Ala	Leu		Gly 285	Ser	Ser	Tyr
60	Lys	Asp 290	Pro	Thr	Ser :	Ser .	Asn . 295	Asp i	Asp	Ser		Ser 300	Asn :	Leu .	Asn	Thr

	Ser 305	Gly	Lys	Thr	Thr	Ser 310	Ser	Thr	Asp	Gly	Pro 315	Lys	Phe	Pro	Gly	Asp 320
5	Lys	Ser	Ser	Thr	Thr 325	Gln	Asn	Asn	Asn	Gln 330	Gln	Lys	Lys	Gly'	Ile 335	Gln
	Val	Leu	Pro	Asp 340	Gly	Arg	Val	Thr	Asn 345	Ile ,	Pro	Gļn	GÌy	Met 350	Val	Thr
10	Asp	Gln	Phe 355	Gly	Меt	Ile	Gly	Leu 360	Leu	Thr	Phe		Arg , 3.65	Ala	Ala	Glu
15	Thr	Asp 370	Pro	Gly	Met	Val	His 375	Leu	Ala '	Leu	Gly	Ser 3 ¹ 80	Ążp	Leu	Thr	Thr
	Leu 385	Gly	Leu	Asn ——		Asn 390		Pro	Glu	Asn	Leu 395	Tyr	Pro	Lys	Phe	Ala 400
20	Ser	Pro	Trp	Ala	Ser 405	Ser	Pro	Суз	Arg	Pro 410	Gln	Asp	Ile	Asp	Phe 415	His
				420				,	425				_	Lys 430		
25			435					440					445	Tyr		
30		450					455					460		Val		
	465			_	_	470	_	ŧ	_		475			Trp		480
35			n		485					490				_	495	Arg
40				500					505					Val 510		
40			515					520					525	His	Leu	Pro
45	Ser	Thr 530	Phe	Asn	Tyr	Asn	Pro 535	Ala	Gln	Gln	Ala	Phe 540	Xaa			
50	(2)							NO: 1								
<i>.</i>			(1). }	()	A) L B) T	ENGT YPE :	H: 5	ERIS 8 am no a lin	ino d		s					
55				SEQ	UENC	E DE	SCRI	PTIO	N: S	_						
	1				5					10				Ile	15	
60	Leu	Cys	Ser	Leu 20	Val	Ile	Gln	Ile	Ser 25	Leu	Lys	Thr	Ile	Arg 30	Asp	Ile

	Thr	Lev	35		Met	Val	Gly	Ile 40		Phe	Ser	Ile	Ser 45		Ser	As
. 5	Lys	: Il∈ 50	Asn	Ile	. Asn	Ser	Arg		Trp	Xaa		. '				
						•	,									
10	(2)	INF	ORMA	,	1						1		1	į		
			(i)	(ENCE (A) I (B) I	ENGI	H: 2	202 a	mino		ds.	•	•			
15			(xi)	((D) I	OPOL	OGY:	lin	ear	EO I	D, NO	: 14	0:			
20	Met 1	Thr	Leu			•								Leu	Leu 15	
20	Leu	Leu	Ser	Ala 20	Ala	Vaľ	Cys	Arg	Ala 25	Glu	Ala	Gly	Leu	Glu 30	Thr	Glı
25	Ser	Pro	Val 35	Arg	Thr	Leu	Gln	Val 40	Glu	Thr	Leu	Val	Glu 45	Pro	Pro	Glu
	Pro	Cys 50	Ala	Ģlu	Pro	Ala	Ala 55	Phe	Gly	Asp	Thr	Leu 60	His	Ile	His	туз
30	Thr 65	Gly	Ser	Leu	Val	Asp 70	Gly	Arg	Ile	Ile	Asp 75	Thr	Ser	Leu	Thr	Arg 80
35	Asp	Pro	Leu	Val	Ile 85	Glu	Leu	Gly	Gln	Lys 90	Gln	Val	Ile	Pro	Gly 95	Let
	Glu	Gln	Ser	Leu 100	Leu	Asp	Met	Cys	Val 105	Gly	Glu	Lys	Arg	Arg 110	Ala	Ile
40	Ile	Pro	Ser 115	His	Leu	Ala	Tyr	Gly 120	Lys	Arg	Gly	Phe	Pro 125	Pro	Ser	Val
		130	Asp	•			135					140				
45	145		Asn			150					155					160
50			Ala		165					170					175	
	Tyr	Arg	Lys	Ala 180	Asn	Arg	Pro	Lys	Val 185	Ser	Lys	Lys	Lys	Leu 190	Lys	Glu
55	Glu	Lys	Arg 195	Asn	Lys	Ser	Lys	Lys 200	Lys	Xaa						

(2) INFORMATION FOR SEQ ID NO: 141: 60

			(i) :	C	A) L B) T	engt Ype :	H: 2	17 a no a	mino cid		ds	, ,	•			
5			(xi)	SEQ		OPOLA E DES				EQ II	D NO	: 14	1:			
	Met 1	Phe	Leu	Arg	Leu 5	Tyr	Leu	Ile	Ala	Arg 10	Val	Met	Leu	Leu	His 15	Ser
10	Lys	Leu	Phe	Thr 20	Asp	Ala	Ser	Ser	Arg 25	Ser	Ile	Gly	Ala '	Leu 30	Asn	Lys
15	Ile	Asn	Phe 35		Thr	Arg i	Phe	Val 40	Met	Lys	Thr	Leu	Met 45	Thr	Ile	Cys
	Pro	Gly 50	Thr	Val	Leu	Leu	Val 55	Phe	Ser'	Ile	Ser	Leu 60	Trp	Ile	Ile	Ala
20	Ala 65	Trp	Thr	Val	Arg	Val 70		Glu	Ser	Pro	Glu 75	Ser	Pro	Ala	Gln	Pro 80
	Ser	Gly	Ser	Ser	Leu 85	Pro	Ala	Trp	Tyr	His 90	Asp	Gln	Gln	Asp	Val 95	Thr
25	Ser	Asn	Phe	Leu 100	Gly	Ala	Met	Trp	Leu 105	Ile	Ser	Ile	Thr	Phe 110	Leu	Ser
30	Ile	Gly	Туг 115	Gly	Asp	Met	Val	Pro 120	His	Thr	Tyr	Cys	Gly 125	Lys	Gly	Val
-	Cys	Leu 130	Leu	Thr	Gly	Ile	Met 135	Gly ,	Ala	Gly	Cys	Thr 140	Ala '	Leu	Val	Val
35	Ala 145	Val	Val	Ala	Arg	Lys 150	Leu	Glu	Leu	Thr	Lys 155	Ala	Glu	Lys	His	Val 160
	His	Asn	Phe	Met	Met 165	Asp	Thr	Gln	Leu	Thr 170	Lys	Arg	Ile	Lys	Asn 175	Ala
40	Ala	Ala	Asn	Val 180	Leu	Arg	Glu	Thr	Trp 185	Leu	Ile	Tyr	Lys	His 190	Thr	Lys
45	Leu	Leu	Lys 195	Lys	Ile	Asp	His	Ala 200	Lys	Val	Arg	Lys	His 205	Gln	Arg	Lys
15	Phe	Leu 210		Ser	Tyr	Pro	Pro 215	Val	Xaa							
50	(2)	INF	ORMA	TION	FOR	SEQ	ID 1	NO:	142:					•		
			(i)	SEQU		CHA ENGI					ds					
55			(xi)	((B) 1 (D) 1	YPE:	ami .OGY :	no a	cid ear): 14	2:			
60	Met 1			_		Val				_	Gln	_		Ser	Asp 15	Ser

	Me	t Va	l Gly	у Туг 20	val	l Leu	ı Gly	/ Pro	25		: Leu	Ile	? Thi	Leu 30		Gly
5	Va.	l Va	1 Va:	l Ala	a Val	Va]	l Met	Tyr 40		Gln	Lys	Lys	Lys 45		v Val	. Asp
10	Arg	J Let 50	ı Arç	g His	His	: Leu	Lev 55	Pro	Met	Tyr	Ser	Тут 60) Pro	Ala	Glu
	Glu 65	Let	ı His	s Glu	Ala	Glu 70	Glm	Glu	Leu	Leu	Ser 75		Met	Gly	Asp	Pro 80
15	Lys	va]	l Val	l His	Gly 85	Trp	Gln	Ser	Gly	Tyr 90	Gln	His	Lys	Arg	Met 95	
	Lev	Leu	ı Asp	Val 100		Thr										
20																
	(2)	INF		TION												1
25				(A) I B) T D) T	ENGT YPE : OPOL	H: 1 ami OGY:	no a	mino cid ear	aci		: 14	3:			
30	Met 1	Arg		Cys										Leu	Pro .15	Phe
35	Ser	Leu	Val	Ser 20	Met	Leu	Val	Thr	Gln 25	Gly	Leu	Vaļ	Tyr	Gln 30	Gly	Тут
	Leu	Ala	Ala 35	Asn	Ser	Arg	Phe	Gly 40	Ser	Leu	Pro	Lys	Val 45	Ala	Leu	Ala
40	Gly	Leu 50	Leu	Gly	Phe	Gly	Leu 55	Gly	Lys	Val	Ser	Tyr 60	Ile	Gly	Val	Cys
	Gln 65	Ser	Lys	Phe	His	Phe 70	Phe	Glu	Asp	Gln	Leu 75	Arg	Gly	Ala	Gly	Phe 80
45	Gly	Pro	Gln	His	Asn 85	Arg	His	Cys	Leu	Leu 90	Thr	Cys	Glu	Glu	Cys 95	Lys
50	Ile	Lys	His	Gly 100	Leu	Ser	Glu	Lys	Gly 105	Asp	Ser	Gln	Pro	Ser 110	Ala	Ser
55	(2)			PION												
			(i) S	SEQUE ()				ERIST O ami								
60								10 ac			-					

```
(D) TOPOLOGY: linear
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 144:
     Met Lys Asn Asp Arg Asn Gln Gly Phe Ser Leu Leu Gln Leu Ile Asp
 5
                                          10
     Trp Asn Lys Pro
10
      (2) INFORMATION FOR SEQ ID NO: 145:
             (i) SEQUENCE CHARACTERISTICS:
15
                (A) LENGTH: 30 amino acids
                    (B) TYPE: amino acid '
                    (D) TOPOLOGY: linear
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 145:
20
     Met Gly Thr Gln Pro Pro Val Val Ala Gly Phe Thr Ile Pro Met Leu
      Gly Tyr Thr Val Arg Val Leu Thr Phe His Leu Ser Cys Ser
                  20
                                      25
25
      (2) INFORMATION FOR SEQ ID NO: 146:
30
             (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 99 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 146:
35
      Met Lys Ile Pro Val Leu Pro Ala Val Val Leu Leu Ser Leu Leu Val
                       5
      Leu His Ser Ala Gln Gly Ala Thr Leu Gly Gly Pro Glu Glu Glu Ser
40
                                      25
      Thr Ile Glu Asn Tyr Ala Ser Arg Pro Glu Ala Phe Asn Thr Pro Phe
45
      Leu Asn Ile Asp Lys Leu Arg Ser Ala Phe Lys Ala Asp Glu Phe Leu
      Asn Trp His Ala Leu Phe Glu Ser Ile Lys Arg Lys Leu Pro Phe Leu
50
      Asn Trp Asp Ala Phe Pro Lys Leu Lys Gly Leu Arg Ser Ala Thr Pro
      Asp Ala Gln
55
      (2) INFORMATION FOR SEQ ID NO: 147:
```

```
(i) SEQUENCE CHARACTERISTICS:
                      (A) LENGTH: 8 amino acids
                      (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
  5
               (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 147:
       Met Val Trp Gly Leu Leu Gly
                        5
 10
       (2) INFORMATION FOR SEQ ID NO: 148:
              (i) SEQUENCE CHARACTERISTICS:
 15
                     (A) LENGTH: 39 amino acids
                     (B) TYPE: amino acid
                     (D) TOPOLOGY: linear
              (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 148:
 20
       Met Leu Pro Leu Leu Ser Leu Leu Phe Leu Phe Phe Ser Thr Val Ser
                   5
                                           10
       Ser Phe Cys Gly Met Pro Leu Arg Ala His Thr Arg Ala Xaa Ala His
                   20
 25
       Thr Arg Thr Phe Ala Ser Arg
               35
 30
       (2) INFORMATION FOR SEQ ID NO: 149:
              (i) SEQUENCE CHARACTERISTICS:
                     (A) LENGTH: 131 amino acids
35
                     (B) TYPE: amino acid
                     (D) TOPOLOGY: linear
              (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 149:
      Met Ile Cys Glu Thr Lys Ala Arg Lys Ser Ser Gly Gln Pro Gly Arg
40
      Leu Pro Pro Pro Thr Leu Ala Pro Pro Gln Pro Pro Leu Pro Glu Thr
45
      Ile Glu Arg Pro Val Gly Thr Gly Ala Met Val Ala Arg Ser Ser Asp
      Leu Pro Tyr Leu Ile Val Gly Val Val Leu Gly Ser Ile Val Leu Ile
                              55
50
      Ile Val Thr Phe Ile Pro Phe Cys Leu Trp Arg Ala Trp Ser Lys Gln
      Lys His Thr Thr Asp Leu Gly Phe Pro Arg Ser Ala Leu Pro Pro Ser
55
                                           90
      Cys Pro Tyr Thr Met Val Pro Leu Gly Gly Leu Pro Gly His Gln Ala
                  100
                                      105
60
     Val Asp Ser Pro Thr Ser Val Ala Ser Val Asp Gly Pro Val Leu Met
```

120 125 115 Gly Ser Thr 130 5 (2) INFORMATION FOR SEQ ID NO: 150: 10 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 32 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 150: 15 Met Gly Ala Pro Ser Leu Thr Met Leu Leu Leu Leu Lys Val Gln Pro 1 5 10 Arg Arg Thr Gln Ala Phe Asp Ala His Trp Val Gly Leu Pro Leu Leu 20 20 25 30 25 (2) INFORMATION FOR SEQ ID NO: 151: (i) SEQUENCE CHARACTERISTICS: 30 (A) LENGTH: 14 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 151: 35 Met Cys Leu Ile Phe Leu Leu Leu Leu Leu Ser Phe Ser 5 10 40 (2) INFORMATION FOR SEQ ID NO: 152: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 8 amino acids (B) TYPE: amino acid 45 (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 152: His Pro His Gln Asp Ser Gln Pro 5 50 (2) INFORMATION FOR SEQ ID NO: 153: 55 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 68 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 153: 60

	Met 1		Thr	Ser	Tyr 5	Ile	Leu	Arg	Leu	Thr 10	Val	. Val	_, Val	Ser	Val 15	Va1
5	Ile	Tyr	Leu	Ala 20		His	Pro	Leu	Leu 25		Phe	Ser	Leu	Glu 30	Ser	Pro
	Leu	Leu	Val 35	Pro	Trp	Arg	Asp	Cys 40		Gln	Asn	Ile	Trp 45	Lys	Ser	Gly
10	Ser	Val 50	Trp	Tyr	Lys	Arg	Trp 55	Thr	Leu	Pro	His	Met 60		Val	Cys	Cys
15	Gln 65	Asp	Leu	His			J			•		,			•	
	(2)	INF	ORMA	rio'n	FOR	SEQ	ID 1	NO: I	154:							
20			(i)	(A) L	ENGT	H: 2	ERIS 6 am no a	ino		s					
25				SEQ	D) T UENC	OPOL E DE	OGY: SCRI	lin PTIO	ear N: S						`	
	1			ì	5					10	Leu	As'n	Leu	Ile	Leu 15	Thr
30	Ser	Ile	Arg	Ile 20	Leu	Glu	Arg	Gln	Asn 25	Met						
35	(2)							No: :				1 1				
			(i) :	(. (:	А) L В) Т	ENGT YPE :	H: 1 ami	ERIS 95 a no a lin	mino cid		ds					
40			(xi)					PTIO		EQ I	D NO	: 15	5:			
	Met 1	Asp	Cys	Glu	Val 5	Asn	Asn	Gly	Ser	Ser 10	Leu	Arg	Asp	Glu	Cys 15	Ile
45	Thr	Asn	Leu	Leu 20	Val	Phe	Gly	Phe	Leu 25	Gln	Ser	Cys	Ser	Asp 30	Asn	Ser
50	Phe	Arg	Arg 35	Glu	Leu	Asp	Ala	Leu 40	Gly	His	Glu	Leu	Pro 45	Val	Leu	Ala
	Pro	Gln 50	Trp	Glu	Gly	Tyr	Asp 55	Glu	Leu	Gln	Thr	Asp 60	Gly	Asn	Arg	Ser
55	Ser 65	His	Ser	Arg	Leu	Gly 70	Arg	Ile	Glu	Ala	Asp 75	Ser	Glu	Ser	Gln	Glu 80
	Asp	Ile	Ile	Arg	Asn 85	Ile	Ala	Arg	His	Leu 90	Ala	Gln	Val	Gly	A sp 95	Ser
60	Met	Asp	Arg	Ser	Ile	Pro	Pro	Gly	Leu	Val	Asn	Gly	Leu	Ala	Leu	Gln

				100					105			,		110		
5	Leu	Arg	Asn 115	Thr	Ser	Arg	Ser	Glu 120	Glu	Asp	Arg	Asn	Arg 125	Asp	Leu	Ala
J	Thr	Ala 130	Leu	Glu	Gln	Leu	Leu 135	Gln	Ala	Tyr	Pro	Arg 140	Asp	Met	Glu	Lys
10	Glu 145	Lys	Thr	Met	Leu	Val 150	Leu	Ala	Leu	Leu	Leu 155	Ala	Lys	Lys	Val	Ala 160
	Ser	His	Thr	Pro	Ser 165	Leu	Leu	Arg	Asp	Val 170	Phe	His	Thr	Thr	Val 175	Asn
15	Phe	Ile	Asn	Gln 180	Asn	Leu	Ärg	Thr	Tyr 185	Val	Arg	Ser	Leu	Ala 190	Arg	Asņ
20	Gly	Met	Asp 195	1		,										
										1						
	(2)	INFO	ORMA!	PION	FOR	SEQ	ID I	NO:	156:				•		- X	
25 .			(i)	1 (A) L B) T	ENGI YPE:	H: 9 ami	1 am	ino cid		s	ı		•		
20			(xi)	SEQ				lin PTIO		EQ I	D NO	: 15	6 :			
30	Met 1	Ser	Leu	Ser	Leu 5	Val	Ser	Val	Ser	Val 10	Gly	Pro		Thr	Leu 15	Ala
35	Cys	Ser	Phe	Leu 20	Arg	Pro	Lys	Ala	Arg 25	Pro	Ser	Lys	Arg	Ser 30	Pro	Arg
	Asn	Tyr	Thr 35	Asp	Ser	Thr	Ser	Pro 40		Gly	Pro	Arg	Ala 45	Pro	Arg	Gly
40	Gly	Ala 50	-	Arg	Leu	Ser	Ser 55		Gln	Asn	Ser	Ser 60		Lys	Gly	Val
45	Ala 65		Ala	Lys	Ala	Ser 70		Arg	Pro	Val	Leu 75		Phe	Leu	Pro	Gly 80
	Pro	Trp	Ser	Ser	Хаа 85		Xaa	Ala	. Phe	Leu 90		!				
50	(2)	INF	ORMA	TION	FOR	SEC	ID	NO:	157 :							
			(i)	SEQU							3 _					
55			/a 1		(B) :	TYPE TOPO	: am:	31 ar ino a : lim	acid near). 1s	: 7 .			
				SEÇ	_										<i>a</i> :	
60	Met 1		Thr	Leu	Sex		Glu	ı Cys	Ser	: Gly 10		Ala	Th	Leu	Gly 15	Leu

Cys Leu Val Val Pro Trp Asn Ser Ser Gly Leu Ser Gln Pro Pro 20 25 5 (2) INFORMATION FOR SEQ ID NO: 158: (i) SEQUENCE CHARACTERISTICS: 10 (A) LENGTH: 91 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 158; 15 Met Lys Phe Leu Ala Val Leu Val Leu Leu Gly Val Ser Ile Phe Leu Val Ser Ala Gln Asn Pro Thr Thr Ala Ala Pro Ala Asp Thr Tyr Pro 20 Ala Thr Gly Pro Ala Asp Asp Glu Ala Pro Asp Ala Glu Thr Thr Ala Ala Ala Thr Thr Ala Thr Thr Ala Ala Pro Thr Thr Ala Thr Thr Ala 25 55 Ala Ser Thr Thr Ala Arg Lys Asp Ile Pro Val Leu Pro Lys Trp Val 30 Gly Asp Leu Pro Asn Gly Arg Val Cys Pro Xaa 35 (2) INFORMATION FOR SEQ ID NO: 159: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 89 amino acids (B) TYPE: amino acid 40 (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 159: Met Ile Ile Ser Leu Phe Ile Tyr Ile Phe Leu Thr Cys Ser Asn Thr 45 Ser Pro Ser Tyr Gln Gly Thr Gln Leu Gly Leu Gly Leu Pro Ser Ala Gln Trp Trp Pro Leu Thr Gly Arg Arg Met Gln Cys Cys Arg Leu Phe 50 Cys Phe Leu Leu Gln Asn Cys Leu Phe Pro Phe Pro Leu His Leu Ile 55 55 Gln His Asp Pro Cys Glu Leu Val Leu Thr Ile Ser Trp Asp Trp Ala Glu Ala Gly Ala Ser Leu Tyr Ser Pro 60

	(2)	INFO	ORMAT	ЙOI	FOR	SEQ	ID N	10: 1	.60 :						ı	
5			(i) S	~ (c	A) LI B) T	engti YPE :	H: 1 ami	ERIST 74 am no ac line	nino cid	_	ds		i			
10			(xi)							EQ IÌ	ONO	: 160	0:			
10	Met 1	Ser	Ser	Ala	Ala 5	Ala	Asp	His	Trp	Ala 10	Trp	Leu	Leu	Val	Leu 15	Ser
15	Phe	Val	Phe	Gly 20	Cys	Asn	Val	Leu	Arg 25	Ile	Leu	Leu	Pro	Ser 30	Phe	Ser
	Ser	Phe	Met 35	Ser	Arg	Val	Leu	Gln 40	Lys	Asp	Ala	Glu	Gln 45	Glu	Ser	Gln
20	Met	Arg 50	Ala	Glu	Ile	Gln	Asp 55	Met	Lys	Gln	Glu	Leu 60	Ser	Thr	Val	Asn
25	Met 65		Asp	Glu	Phe	Ala 70	Arg	Tyr	Ala	Arg	Leu 75	Glu	Arg	Lys	Ile	Asn 80
25	Lys	Met	Thr	Asp	Lys 85	Leu	Lys	Thr	His	Val 90	Lys	Ala	Arg	Thr	Ala 95	Gln
30	Leu	Ala	Lys	Ile 100	Lys	Trp	Val	Ile	Ser 105	Val	Ala	Phe	Tyr	Val 110	Leu	Gln
	Ala	Ala	Leu 115	Met	Ile	Ser	Leu	Ile 120	Trp	Lys	Tyr	Tyr	Ser 125	Val	Pro	Val
35	Ala	Val 130	Val	Pro	Ser	Lys	Trp 135		Thr	Pro	Leu	Asp 140	Arg	Leu	Val	Ala
40	Phe 145		Thr	Arg	Val	Ala 150		Gly	Val	Gly	Ile 155	Thr	Cys	Trp	Ile	Leu 160
40	Val	. Cys	Asn	Lys	Val 165		Ala	Ile	Val	Leu 170		Pro	Phe	Ser		
45	(2)	INF	ORMA	TION	FOR	SEQ	ID	NO:	161:							
50				1	(A) I (B) I (D) I	LENGI TYPE : TOPOI	TH: 4 am:	TERIS 15 an ino a ino teritorical	nino acid near	ació		· 16	:1.			:
55		: Gly	(XI)			. Asn					Lys			Leu	Leu 15	
	Let	ı Val	l Gln	Cys 20		ı Asn	сув	: Cys	Arg		: Asn	. Met	. Leu	тут 30		ı Ile
60	D)-		, yen	. 714	. wie	, her	. τ1 <i>-</i>	. Wie	Tage	. Dhe	Ser	· Acr	, Hic			

35 40 45 . 2 (2) INFORMATION FOR SEQ ID NO: 162: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 23 amino acids (B) TYPE: amino acid 10 (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 162: Met Val Ala Ser Thr Leu Val Thr Asn Leu Phe Gly Val Ala Phe Ala 15 Thr Thr Ala Ala Thr Arg Ala 20 20 (2) INFORMATION FOR SEQ' ID NO: 163: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 70 amino acids 25 (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 163: Met Leu Met Ala Pro Val Val Cys Leu Ser Phe Ser Pro Cys Pro Ala 30 Asp Thr Ser Leu Thr Gly Asp Gly Leu Lys Ala Gly Leu Glu Arg Gly 25 35 Xaa Ala Leu Val Thr Leu Phe Asp Ser Val Thr His Phe Leu Ala His 40 Thr Leu Phe Glu Leu Leu Asp Phe Gln Leu Ala Phe Leu Arg Ser Gly 55 40 Lys Gln Thr Ala Pro His 45 (2) INFORMATION FOR SEQ ID NO: 164: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 323 amino acids 50 (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 164: Met Leu Leu Leu Leu Leu Gly Ser Gly Gln Gly Pro Gln Gln 55 5 Val Gly Ala Gly Gln Thr Phe Glu Tyr Leu Lys Arg Glu His Ser Leu

Ser Lys Pro Tyr Gln Gly Val Gly Thr Gly Ser Ser Ser Leu Trp Asn

			35	•				40		•			45			
_	Leu	Met 50	Gly	Asņ	Ala	Met	Val 55	Met	Thr	Gln	Tyr	Ile 60	Arg	Leu	Thr	Pro
5	Asp 65	Met	Gln	Ser	Lys	Gln 70	Gly	Ala	Leu	Trp	Asn 75	Arg ,	Val	Pro	Cys	Phe 80
10	Leu	Arg	Asp	Trp	Glu 85	Leu	Gln	Val	His	Phe 90	'Lys	Ile	His	Gly	Gln 95	Gly
	Lys	Lys	Asn	Leu 100	His	Gly	Asp	Gly	Leu 105	Ala	Ile	Trp	Tyr	Thr 110	Arg	Asn
15	Arg	Met	Gln 115	Pro	Gly	Pro	Val	Phe 120	Gly	Asn	Met	Asp	Lys 125	Phe	Val	Gly
20	Leu ·	Gly 130	Val	Phe	Val	Asp	Thr 135		Pro	Asn	Glu	Glu 140		Gln	Gln	Glu
20	Arg 145		. Phe	Pro	Tyr	Ile 150		Ala	Met	Val	Asn 155	Asn	Gly	Ser	Leu	Ser 160
25	Tyr	Asp	His	Glu	Arg 165		Gly	Arg	Pro	Thr 170		Leu	Gly	Gly	Cys 175	Thr
	Ala	Ile	e Val	Arg 180		Leu	His	Tyr	Asp 185		Phe	Leu	Val	Ile 190	Arg	Тут
30	Val	. Lys	195		Leu	Thr	Ile	Met 200		Asp) Ile	Asp	Gly 205		His	Glu
35	Tr	210	g Asp O	суз	Ile	: Glu	Val 215		Gly	Va]	Arg	220		Arg	Gly	Туз
	Туз 225		e Gly	Thr	: Ser	230		e Thr	Gly	As <u>r</u>	235		Asp	Asn	His	As ₁ 240
40	Val	Il	e Sex	. Leu	Lys 245		ı Phe	e Glu	ı Lev	250		. Glı	ı Arg	Thr	255	Gl:
	Glı	ı Gl	u Ly:	260		Arg	y Ası	o Val	265		ı Pro	Sei	r Val	270		n Me
45	Ly	s Le	u Pro 27!		ı Met	: Thi	r Ala	280		ı Pr	o Pro	Let	289		/ Let	ı Al
50	Le	u Ph 29	e Le 0	ı Ile	e Vai	l Pho	e Ph 29		r Lei	ı Va	l Phe	30		l Phe	e Ala	a Il
	Va 30		e Gl	y Il	e Ilo	e Le		r Ası	n Ly:	s Tr	p Gli 31		u Gl	n Sei	r Ar	32
55	Ar	g Ph	е Ту	r												
	(2	:) II	IFORM	ATIO	n fo	R SE	Q II	NO:	165	:						

		(i)) Seq	QUEN	Œ CI	HARA	CTER	ISTI	cs:			•			
				(A)	LEN	GTH:	321	ami	no a	cids		!			
							mino							t	
· 5		(xi	i) si	(U) (U)	TOP	OLOG	Y: 1	inea	r		_ , '				
		(323	., 51	-ZOIH	VCE I)ESCI	RIPT	LON:	SEQ	ID 1	NO:	165:			
	Met Pr	ro Se	r Gl	u Ty	r Th 5	ur Ty	⁄r Va	ıl Ly	ys Le	eu Ar l0	g Se	er As	ip C <u>)</u>		
10						•									15
10	Pro Se	er Le	u G1 2	n Tr O :	р Ту	r Th	r Ar	g Al	la G] 25	n Se	r Ly	's Me		g Ai	rg Pr
	Ser Le	u Le	u Le	u Ly	s As	p Il	e Le	u Ly	rs Cy	s Th	r Le	u Le	u Va	l Pł	ne Gl
15							4	U		•		4	5		
	Val Tr 5	p Ilo	e Le	u Ty	r Il	ei Le 5	u Ly 5	s Le	u As	п Ту	r Th	r Th 0	r Gl	u Gl	u Cys
	Asp Me	+ T		 - \ 7-1		_				٠.					
20	Asp Me 65	c ny:	з г.Х:	s Mei	7 H1	s Ty: 0 '	r Va	l As	p Pr	o As 7	рНi 5	s Va	l Ly	s Ar	g Ala 80
	Gln Ly	ѕ Тут	Ala	a Glr 85	Glr	ı Vai	l Let	ı Gl	n Ly: 9	s Gl	и Су	s Arg	g Pr	o Ly 9	
25	Ala Ly:	s Thr	Ser 100	Met	Ala	Let	ı Let	1 Pho	e Gl: 5	ı His	a Ar	Туз	Sei 110	r Va	
30	Leu Le	1 Pro 115	Phe	val	Glr	Lys	120	Pro) Lys	a Ası	Sei	Glu 125		a Gl	u Ser
	Lys Tyr 130	Asp	Pro	Pro	Phe	Gly 135	Phe	: Arg	J Lys	Phe	Ser 140	: Ser	Lys	Va]	l Gln
35	Thr Leu 145	Leu	Glu	Leu	Leu 150	Pro	Glu	His	asp	Leu 155	Pro	Glu	His	Leu	1 Lys 160
	Ala Lys	Thr	Cys	Arg 165	Arg	Cys	Val	Val	Ile 170	Gly	Ser	Gly	Gly	Ile 175	
40	His Gly	Leu	Glu 180	Leu	Gly	His	Thr	Leu 185	Asn	Gln	Phe	Asp	Val 190	Val	Ile
45	Arg Leu	Asn 195	Ser	Ala	Pro	Val	Glu 200	Gly	Tyr	Ser	Glu	His 205	Val	Gly	Asn
	Lys Thr 210	Thr	Ile	Arg	Met	Thr 215	Tyr	Pro	Glu	Gly	Ala 220	Pro	Leu	Ser	Asp
50	Leu Glu 225	Tyr	Tyr	Ser	Asn 230	Asp	Leu	Phe	Val	Ala 235	Val	Leu	Phe	Lys	Ser 240
	Val Asp	Phe	Asn	Trp 245	Leu	Gln	Ala	Met	Val 250	Lys	Lys	Glu	Thr	Leu 255	Pro
55	Phe Trp	Val	Arg 260	Leu	Phe	Phe	Trp	Lys 265	Gln	Val	Ala	Glu	Lys 270	Ile	Pro
60	Leu Gln	Pro : 275	Lys	His	Phe	Arg	Ile 280	Leu	Asn	Pro	Val	Ile 285	Ile	Lys	Glu

```
Thr Ala Phe Xaa His Pro Ser Val Leu Arg Ala Ser Val Lys Val Leu
                             295
     Gly Ala Glu Ile Arg Thr Ser Pro Gln Ser Val Ser Leu Pro Leu Ser
 5
                         310
                                             315
     Xaa
10
      (2) INFORMATION FOR SEQ ID NO: 166:
             (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 31 amino acids
15
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 166:
20
     Met Thr Leu Asp Val Gln Thr Val Val Val Phe Ala Val Ile Val Val
                       5
                                      1 10
      Leu Leu Leu Val Asn Val Ile Leu Met Phe Phe Leu Gly Thr Arg
                  20
                                      25
25
      (2) INFORMATION FOR SEQ ID NO: 167:
30
             (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 72 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 167:
35
      Met Leu Pro Leu Leu Phe Cys Ala Phe Cys Leu His Lys Leu Gly Pro
                       5
      Leu Leu Phe Leu Tyr Asp Val Leu Met Xaa His Glu Ala Val Met Arg
40
                                       25
      Thr His Gln Ile Gln Leu Pro Asp Pro Glu Phe Pro Ser Gln Gln Asn
45
      Gln Val Leu Asn Lys Thr Leu Phe Asn Lys Leu Lys Lys Lys Lys
      Lys Lys Lys Xaa Xaa Xaa Lys Lys
50
      (2) INFORMATION FOR SEQ ID NO: 168:
55
             (i) SEQUENCE CHARACTERISTICS:
                     (A) LENGTH: 282 amino acids
                     (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 168:
60
```

	Met 1	Ala	a Ser	Arg	r Gly 5	Arg	Arg	Pro	Glu	His		Gly	Pro	Pro	Glu 15	
5	Phe	Тут	: Asp	Glu 20	Thr	Glu	Ala	Arg	Lys 25		Val	Arg	Asn	Ser 30	!	Met
	Ile	Asp	Ile 35	Gln	Thr	Arg	Met	Ala 40		Arg	Ala	Leu	Glų 45	Leu	Leu	Tyr
10	Leu	Pro 50	Glu	Asn	Lys	Pro	Cys 55	Tyr	Leu	Leu	Asp	Ile 60	Gly	Cys	Gly	Thr
15	Gly 65	Leu	Ser	Gly	Ser	Туг 70	Leu	Ser	Asp	Glu	Gly 75	His	Tyr	Trp	Val	Gly 80
	Leu	Asp	Ile	Ser	Pro 85	Ala	Met	Leu	Asp	Glu 90	Ala	Val	Asp	Arg	Glu 95	Ile
20	Glu	Gly	Asp	Leu 100	Leu	Leu	Gly	Asp	Met 105	Gly	Gln	Gly	Ile	Pro 110	Phe	Lys
	Pro	Gly	Thr 115	Phe	Asp	Gly	Суз	Ile 120	Ser	Ile	Ser	Ala	Val 125	Gln	Trp	Leu
25	Cys	Asn 130	Ala	Asn	Lys	Lys	Ser 135	Glu /	Asn	Pro	Ala	Lys 140	Arg	Leu	Tyr	Cys
30	Phe 145	Phe	Ala	Ser	Leu	Phe 150	Ser	Val	Leu	Val	Arg 155	Gly	Ser	Arg	Ala	Val 160
	Leu	Gln	Leu	Tyr	Pro 165	Glu	Asn	Ser	Glu	Gln 170	Leu	Glu	Leu	Ile	Thr 175	Thr
35	Gln	Ala	Thr	Lys 180	Ala	Gly	Phe	Ser	Gly 185	Gly	Met	Val	Val	Asp 190	Tyr	Pro
	Asn	Ser	Ala 195	Lys	Ala	Lys	Lys	Phe 200	Tyr	Leu	Cys	Leu	Phe 205	Ser	Gly	Pro
1 0		210		Ile			215					220				
15	Pro 225	Arg	Glu	Ser		Phe 230	Thr	Asn	Glu	Arg	Phe 235	Pro	Leu	Arg	Met	Ser 240
	Arg				245					250					255	
50	Glu	Arg	His	Arg 260	Arg	Gln (Gly		Glu 265	Val	Arg	Pro .		Thr 270	Gln	Tyr
· Æ	Thr	Gly	Arg 275	Lys .	Arg	Lys :		Arg 280	Phe :	Xaa						
55																

(2) INFORMATION FOR SEQ ID NO: 169:

60

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 amino acids

		•	(xi)	((D) 1	TYPE: TOPOI E DE	.OGY :	lir	ear	EQ I	D NC): 16	i9:			
5	Met 1		Gly	Lys	Thr 5		Phe	Gln	Ser	Tyr 10	Lys	Ser	Phe	Ser	Arg 15	Lys
10	Leu	Met	Val	Cys 20	Pro	Ser	Thr			,		•				
15	(2)	INF	ORMA' (i) (xi)	SEQU))	ENCE (A) I (B) I	CHA ENGI YPE:	RACT H: 3 ami OGY:	ERIS 28 a no a lin	TICS mino cid ear	: aci EQ I		i . 17	0.			
20	Met 1							•						Arg	His 15	Gly
25	Ala	Gln	Gly	Lys 20	Pro	Ser	Pro	Asp	Ala 25	Gly	Pro	His	Gly	Gln 30	Gly	Arg
	Val	His	Gln 35	Ala	Ala	Pro	Leu	Ser 40	Asp	Ala	Pro	His	Asp 45	Asp	Ala	His
30		50					55					60			Val	
35	65					70					75				Leu	80
					85					90					Trp 95 Arg	
40				100					105					110	Asp	
45		Gly	115				Glu	120				Ala	125		Gly	
	Tyr 145	130 Ala	Pro	Gly	Glu	Glu 150	135 Phe	His	Asp	Val	Glu 155	140 Asp	Ala	Glu	Thr	Туг 160
50	Lys	Lys	Met	Leu	Ala 165	Arg	Asp	Glu	Arg	Arg 170	Phe	Arg	Val	Ala	Asp 175	
55	Asp	Gly	Asp	Ser 180	Met	Ala	Thr	Arg	Glu 185	Glu	Leu	Thr	Ala	Phe 190	Leu	His
	Pro	Glu	Glu 195	Phe	Pro	His	Met	Arg 200	Asp	Ile	Val	Ile	Ala 205	Glu	Thr	Leu
60	Glu	Asp	Leu	qzA	Arg	Asn	Lys	Asp	Gly	Tyr	Val	Gln	Val	Glu	Glu	Tyr

		210					215					220	ı			
5	Ile 225	Ala	Asp	Leu	туг	Ser 230	Ala	Glu	Pro	Gly	Glu 235		Glu	Pro	Ala	Trp 240
	Val	Gln	Thr	Glu	Arg 245	Gln	Gln	Phe	Arg	Asp 250	Phe	Arg	Asp	Leu	Asn 255	Lys
10	Asp	Gly		Leu 260	Asp '	Gly	Ser	Glu	Val 265	Gly	His	Trp	Val	Leu 270	Pro	Pro
	Ala	Gln	Asp 275	Gln	Pro	Leu	Val	Glu 280	Ala	Asn	His	Leu	Leu 285	His	Glu	Ser
15	Asp	Thr 290	Asp	Lys	Asp		Arg 295	Leu	Ser	Lys	Ala	Xaa 300	Ile	Leu	Gly	Asn
20	Trp 305	Asn	Met	Phe	Val	Gly 310	Ser	Gln	Ala	Thr	Asn 315	Tyr	Gly	Glu	Asp	Leu 320
	Thr	Arg	His	His	Asp 325	Glu	Leu	Xaa		1	•					
25	(2)	INF	ORMA!	rion	FOR	SEQ	ID 1	vo: 1	L 71 ;			1			``	
30			(i) (xi)	() ()	A) L B) T D) T	engt YPE : OPOL	H: 6 ami OGY:	ERIS 9 am no a lin PTIO	ino d cid ear	acid		: _, 17:	1:			
35	Met 1	Cys	Trp	Leu	Arg 5	Ala	Trp	Xaa	Gln	Ile 10	Xaa	Leu	Pro	Val	Phe 15	Xaa
	Ser	Xaa	Phe	Leu 20	Ile	Gln	Leu	Leu	Ile 25	Ser	Phe	Ser	Glu	Asn 30	Gly	Phe
40	Ile	His	Ser 35	Pro	Arg	Asn	Asn	Gln 40	Lys	Pro	Arg	Asp	Gly 45	Asn	Xaa	Glu
1 5		50	Ala Met			Lys	Ser 55	Cys	Gln	Leu	Cys	Thr 60	Glu	Asp	Lys	Lys
50	(2)	INFO	ORMA!	rion	FOR	SEQ	ID I	¥O: 1	172:							
55			(i) : (xi)	(A) L B) T D) T	ENGT YPE : OPOL	H: 1 ami OGY:	ERIS 60 a no a lin PTIO	mino cid ear	aci		: 17:	2:			
50	Met 1	Trp	Leu	Phe	Ile 5	Leu	Leu	Ser	Leu	Ala 10	Leu	Ile	Ser	Asp	Ala 15	Met

•	Val	Met	Asp	Glu 20	Lys	Val	Lys	Arg	Ser 25	Phe	Val	Leu ,	Asp	Thr	Ala	Ser
5	Ala	Ile	Cys 35	Asn	Tyr	Asn	Ala	His 40	Tyr	Lys	Asn	His	Pro 45	Lys	Tyr	Trp
10	Cys	Arg 50	Gly	Tyr	Phe	Arg	Asp 55	Туг	Суѕ	Asn		Ile 60	Ala	Phe	Ser	Pro
	Asn 65	Ser	Thr	Asn	His	Val 70	Ala	Leu	Lys	Asp	Thr 75	Gly	Asn	Gln	Leu	Ile 80
15	Val	Thr	Met	Ser	Суз 85	Leú	Asn 	Lys	Glu '	Asp 90	Thr	Gly	Trp	Tyr	Trp 95	Cys
	Gly	Ile	Gln	Arg 100	Asp	Phe	Ala	Arg	Asp 105	Asp	Met,	Asp	Phe	Thr 110	Glu	Leu
20	Ile	Val	Thr 115	Asp	Asp	Lys	Gly	Thr 120	Trp	Pro	Met	Thr	Leu 125	Val	Trp	Glu
25	Arg	Leu 130		Gly	Thr	Lys	Pro 135	Glu	Ala	Ala	Arg	Leu 140	Pro	Lys	Leú	Ser
	Ala 145	Arg	Leu	Thr	Ala	Pro 150	Gly	Arg	Pro	Phe	Ser 155	Ser	Phe	Ala	Tyr	Xaa 160
30													•			
	(2)	INF	ORMA	TION	FOR	SEQ	ID 1	NO:	173:			1	i			
35			(i)		ENCE (A) I (B) T	ENGI	н: 1	.23 a	mino		.ds					
40			(xi)	SEQ	D) I					EQ I	D NC	: 17	3:			
	Met 1		Xaa	His	Phe 5	Leu	Leu	Val	Ala	Leu 10	Gln	Ser	Val	Pro	His 15	Cys
45	Pro	His	Leu	Leu 20		Glu	Glu	His	Lys 25	Leu	Суз	Lys	Val	Ser 30	His	Phe
50	Ser	Gly	Val 35		Leu	Val	Thr	Ser 40	-	Gln	Asp	Ser	Ser 45		Tyr	Val
50	Pro	Val 50		Thr	Leu	Phe	Ile 55		Leu	Gly	Pro	Trp 60		Trp	Asp	Leu
55	Хаа 65		Cys	Thr	Ala	Glu 70		Pro	Glu	Ala	Glu 75		Ser	Leu	Arg	Leu 80
	Cys	His	Ser	His	Leu 85		Arg	Xaa	. Asn	Val 90		Pro	Ser	Gln	Ala 95	Ala
60	Glü	Gly	Хаа	Xaa	Xaa	Arg	Gly	Cys	Gln	His	Arg	Gly	Ser	Arg	Glu	Leu

				100					105					110		
5	Thr	Phe	Leu 115	Ser	Ala	Glu	Asn	Glu 120	Ala	Gly	Ile				ì	
	(2)	INF	ORMA:	rion	FOR	SEQ	ID I	NO: :	174:	,		,	٠			
10			(i) :	(A) L B) T	ENGI YPE:	H: 1 ami	ERIS 29 a no a	mino cid		ds					
15			(xi)					lin PTIO		EQ I	D NO	: 17	4:			
13	Met 1	Lys	Val	Gly	Ala 5	Arg	Ile	Arg	Val	Lys 10	Met	Ser	Val	Asn	Lys 15	Ala
20	His	Pro	Val	Val 20	Ser	Thr	His	Trp	Arg 25	Trp	Pro	Ala	Glu	Trp 30	Pro	Glr
	Met	Phe	Leu 35	His	Leu	Ala	Gln	Glu 40	Pro	Arg	Thr	Glu	Val 45	Lys	Ser	Arg
25	Pro	Leu 50	Gly	Leu	Ala	Gly	Phe 55	Ile	Arg	Gln	Asp	Ser 60	Lys	Thr	Arg	Lys
30	Pro 65	Leu	Glu	Gln	Glu	Thr 70	Ile	Met	Ser	Ala	Ala 75	Asp	Thr	Ala	Leu	Trp 80
	Pro	Тут	Gly	His	Gly 85	Asn	Arg	Glu	His	Gln 90	Glu	Asn	Glu	Leu	Gln 95	Lys
35	Tyr	Leu	Gln	Туг 100	Lys	Asp	Met	His	Leu 105	Leu	Asp	Ser	Gly	Gln 110	Ser	Leu
	Gly	His	Thr 115	His	Thr	Leu	Gln	Gly 120	Ser	His	Asn	Leu	Thr 125	Ala	Leu	Asn
40	Ile														-	
45	(2)	INFO	ORMAI	TON	FOR	SEQ	ID 1	vo: 1	L75:							
50				() () ()	A) Li B) T D) T	ENGT YPE: OPOL	H: 3' ami OGY:	ERIST 72 at no a line PTION	mino cid ear	aci		: 17	5:			
	Met													His	Ala	Tro
55	1				5					10			•		15	- 45
	Asn	Lys	Asp	Arg 20	Thr	Gln	Ile	Ala	Ile 25	Cys	Pro	Asn	Asn	His 30	Glu	Val
60	His	Ile	Тут 35	Glu	Lys	Ser	Gly	Ala 40	Lys	Trp	Thr	Lys	Val 45	His	Glu	Leu

	Lys	Glu 50	His	Asn	Gly	Gln	Val 55	Thr	Gly	Ile	Asp	Trp 60	Ala	Pro	Glu	Ser
5	Asn 65	Arg	Ile	Val	Thr	Cys 70	Gly	Thr	Asp	Arg	Asn 75	Ala	Tyr	Val	Trp	Thr 80
10	Leu	Lys	Gly	Arg	Thr 85	Trp	Lys	Pro	Thr	Leu 90	Val	Ile	Leu	Arg ,	Ile 95	Asn
	Arg	Ala	Ala	Arg 100	Cys	Val	Arg	Trp	Ala 105	Pro	Asn	Glu ,	^l Asn	Lys 110	Phe	Ala
15	Val	Gly	Ser 115	Gly	Ser		Val	Ile 120	Ser	Ile	Cys	Tyr	Phe 125	Glu	Gln	Glu
	Asn	Asp 130	Trp	·Trp	Val	Cys	Lys 135	His	Ile	Lys	Lys	Pro 140	Ile	Arg	Ser	Thr
20	Val 145	Leu	Ser	Leu	Asp	Trp 150	His	Pro	Asn	Asn '	Val 155	Leu	Leu	Ala	Ala	Gly 160
25	Ser	Cys	Asp	Phe	Lys 165	Cys	Arg	Ile	Phe	Ser 170	Ala	Tyr	Île	Lys	Gľu 175	Val
	Glu	Glu	Arg	Pro 180	Ala	Pro	Thr	Pro	Trp 185	Gly	Ser	ГÀЗ		Pro 190	Phe	Gly
30			195					Ser 200					205			
	Cys	Phe 210	Ser	Ala	Ser	Gly	Ser 215	Arg	Val	Ala	Trp	Val 220	Ser	His	Asp	Ser
35	Thr 225	Val	Cys	Leu	Ala	Asp 230	Ala	Asp	Lys	Lys	Met 235	Ala	Val	Ala	Thr	Leu 240
40	Ala	Ser	Glu	Thr	Leu 245	Pro	Leu	Leu	Ala	Leu 250	Thr	Phe	Ile	Thr	Asp 255	Asn
				260				Asp	265					270		
45			275					Ser 280					285			
	Lys	Gln 290	Ser	Ser	Gln	Arg	Gly 295	Leu	Thr	Ala	Arg	Glu 300		Phe	Gln	Asn
50	305	-	_	-		310		Glu			315					320
55	Leu	Asp	Ser	Leu	His 325		Asn	Ser	Val	Ser 330		Ile	Ser	Val	Leu 335	
	Gly	Gly	Lys	Ala 340		Cys	Ser	Gln	Phe 345		Thr	Thr	Gly	Met 350		Gly
60	Gly	Met	Ser 355		Trp	Asp	Val	Lys 360		Leu	Glu	Ser	Ala 365		Lys	Asp

Leu Lys Ile Lys 370

J																
	(2)	INF	ORMA	TION	FOR	SEQ	ID I	NO:	176:	,			,			
10				(A) I B) T D) T	ENGI YPE : OPOL	H: 2 ami OGY:	ERIS 16 a no a lin PTIO	mino .cid .ear	aci		: 17	 6: ,			
15	Met 1	Trp	Ser	Ile	Gly 5	Ala	Gly	Ala	Leu	Gly 10	Ala	Ala	Ala	Leu	Ala 15	Leu
20	Leu	Leu	Ala	Asn 20	Thr	Asp	Val	Phe	Leu 25	Ser	Lys	Pro	Gln	Lys 30	Ala	Ala
	Leu	Glu	Тут 35	Leu	Glu	Ąsp	Ile	Asp 40	Leu	Lys	Thr	Leu	Glu 45	Lys	Glu	Pro
25	Arg	Thr 50	Phe	Lys	Ala	Lys	Glu 55	Leu	Trp	Glu	Lys	Asn 60	Gly	Ala	Val	Ile
	Met 65	Ala	Val	Arg	Arg	Pro 70	Gly	Суз	Phe	Leu	Су s 75	Arg	Glu	Glu	Ala	Ala 80
30	Asp	Leu	Ser	Ser	Leu 85	Lys	Ser	Met	Leu	Asp 90	Gln	Leu	Gly	Val	Pro 95	Leu
35	Tyr	Ala	Val	Val 100	Lys	Glu	His	Įle	Arg 105	Thr	Glu	Val	Lys	Asp 110	Phe	Gln
	Pro	Tyr	Phe 115	Lys	Gly	Glu	Ile	Phe 120	Leu	Asp	Glu	Lys	Lys 125	Lys	Phe	Tyr
10	Gly	Pro 130	Gln	Arg	Arg	Lys	Met 135	Met	Phe	Met	Gly	Phe 140	Ile	Arg	Leu	Gly
	Val 145	Trp	Tyr	Asn	Phe	Phe 150	Arg	Ala	Trp	Asn	Gly 155	Gly	Phe	Ser	Gly	Asn 160
15	Leu	Glu	Gly	Glu	Gly 165	Phe	Ile	Leu	Gly	Gly 170	Val	Phe	Val	Val	Gly 175	Ser
50	Gly	Lys	Gln	Gly 180	Ile	Leu	Leu	Glu	His 185	Arg	Glu	Lys	Glu	Phe 190	Gly	Asp
	Lys	Val	Asn 195	Leu	Leu	Ser	Val	Leu 200	Glu	Ala	Ala	Lys	Met 205	Ile	Lys	Pro
55	Gln	Thr 210	Leu	Ala	Ser	Glu	Lys 215	Lys								

(2) INFORMATION FOR SEQ ID NO: 177:

	(i) SEQUENCE CHARACTERISTICS:
	(A) LENGTH: 55 amino acids (B) TYPE: amino acid
	(D) TOPOLOGY: linear
5	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 177:
	Met Lys Pro Val Ser Arg Arg Thr Leu Asp Trp Ile Tyr Ser Val Leu 1 5 10 15
10	Leu Leu Ala Ile Val Leu Ile Ser Trp Gly Cys Ile Ile Tyr Ala Ser 20 25 30
	Met Val Ser Ala Arg Arg Gln Leu Arg Lys Lys Tyr Pro Asp Lys Ile 35 40, 45
15	Phe Gly Thr Asn Glu Asn Leu 50 55
20	
	(2) INFORMATION FOR SEQ ID NO: 178:
25	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 23 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 178:
30	Met Ala Ala Asn Thr Phe Val Leu Ile Met Gly Ile Pro Thr Ser Ala 1 5 10 15
	Asn Ala Xaa Arg Asp Leu Phe
35	
	(2) INFORMATION FOR SEQ ID NO: 179:
40	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 103 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 179:
45	Met Ser Ile Cys His Arg Gly Thr Gly Ile Ala Leu Ser Ala Gly Val 1 5 10 15
50	Ser Leu Phe Gly Met Ser Ala Leu Leu Leu Pro Gly Asn Phe Glu Ser 20 25 30
50	Tyr Leu Glu Leu Val Lys Ser Leu Cys Leu Gly Pro Ala Leu Ile His 35 40 45
55	Thr Ala Lys Phe Ala Leu Val Phe Pro Leu Met Tyr His Thr Trp Asn 50 55 60
	Gly Ile Arg His Leu Met Trp Asp Leu Gly Lys Gly Leu Lys Ile Pro 65 70 75 80
60	Clarkey Tor Clarker Gly Val Val Val Leu Val Leu Thr Val Leu Ser

					85	5				90)		•		9!	5
· 5	Sei	: Mei	t Gly	7 Let 100		A Ala	a Met	E				. ,				
10	(2)	IN	FORMA	SEQU	JENCI	CHZ	ARAC.	NO: TERIS	STICS	S:	ls.					
15				SEC	(B) ((D) (QUENC	IYPE IOPOI E DE	: am LOGY ESCR	ino a : lia IPTIC	acid near ON: S	SEQ I	ID NO					
	Met 1	Thr	Lys	Ala	Ser 5	Ser	Leu	Trp	Pro	Leu 10		Thr	Thr	Cys	Glr: 15	
20	Ser	Gly	Thr	Val 20	Phe	Phe	Phe	: Leu	Phe 25		Phe	Ser	: Cys	Phe 30		Me:
	Gln	Ala	Gln 35	Cys	Asp	Lys	Phe	• Val 40		Trp	Asp	Phe	Phe 45		Phe	Let
25				,									ı			
30	(2)	INF	ORMA													
35				((A) I B) T D) T	ENGT YPE: OPOL	H: 9 ami OGY:	ERIS 6 am no a lin PTIO	ino cid ear	acid		: 18	1:			
40	Met 1	Arg	Arg	Ala	Leu 5	Ile	Pro	Pro	Cys	Arg 10	Gly	Gly	Pro	Ser	Ala 15	Ser
	Asp	Xaa	Cys	Суs 20	Ser	Cys	Ser	Pro	Ser 25	Gly	Phe	Ser	Ala	Gly 30	Arg	Gly
1 5	Arg	Суѕ	Pro 35	Val	Gln	Gly	Cys	Leu 40	Arg	Pro	His	Arg	Val 45	Gln	Leu	Leu
	Arg	Arg 50	Trp	Gly	Pro	Gly	Ser 55	Pro	Ala	Gly	Gln	Arg 60	Leu	Ser	Lys	Gly
50	Phe 65	Gln	Leu	Leu	Arg	Trp 70	Trp	Gly	Pro	Gly	Ser 75	Pro	Ala	Pro	Glu	Pro 80
55	Arg	Lys	Gly	Pro	Phe 85	Pro	Pro	Pro	Asp	Pro 90	Pro	Trp	Pro	Val	Thr 95	Leu

	(2) INFORMATION FOR SEQ ID NO: 162:	
5	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 95 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 182:	
10	Met Leu Glu Thr Thr Lys His Val Gln Ile Ala Cys Met Leu Leu L 1 15 15 10 15	eu
	Thr Cys Gln Ile Phe Leu Pro Ser Ser Leu Ser Pro Ser Phe Ile H 20 25 30	is
15	Ser Leu Thr Asp Ser Phe Ile Pro Leu Lys Lys Leu Tyr Val Cys Pi 35 40 45	he
20	Val Gln Ser Thr Leu Leu Lys Ala Ala Gly Tyr Lys Ser Ile Ser G 50 55 60	lt
	Ala Leu Gly Phe Asp Xaa Leu Leu Cys Ser Ser Ala Arg Phe Val T. 65 70 75	rţ
25	Ile Cys His Thr Tyr Ser Arg Pro Leu Val Thr Cys Ala Leu His 85 90 95	
30	(2) INFORMATION FOR SEQ ID NO: 183: (i) SEQUENCE CHARACTERISTICS:	
35	(A) LENGTH: 27 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 183:	
	Met Ser Val Ile Gly Gly Leu Leu Leu Val Val Ala Leu Gly Pro G 1 5 10 15	ly
40	Gly Val Ser Met Asp Glu Lys Lys Lys Glu Trp 20 25	
45	(2) INFORMATION FOR SEQ ID NO: 184:	
50	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 11 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 184:	
55	Met Ser Gly Gly Leu Ser Phe Leu Leu Leu Val 1 5 10	
	(2) INFORMATION FOR SEQ ID NO: 185:	
60	(i) SEQUENCE CHARACTERISTICS:	

						ENGT YPE:				acid	s					
				(D) T	OPOL	OGY:	lin	ear				_			
5			(xi)	SEQ	UENC	E DES	SCRI)	PTIO	N: S	EQ I	O NO	: 18	5:	ŀ		
	Met 1	Phe	Ala	Asp	Phe 5	Ile	Val	Val	Thr	Ala 10	Thr	Val ,	Gln	Arg	Cys 15	Pro
10	Gly	Ser	Pro	Pro 20	Leu	Ser	Glu	Ile	Leu 25	Tŗp	Lys	Asp	Glu	Pro 30	Phe	Ala
	Ile	Ser	Ser 35	His	Ala	Gly	Leu	Pro 40	Trp	Leu	Ser	Ser	Trp ,45	Pro	Ala	Pro
15	Pro	Trp 50	Thr	Trp	Ser	Trp	Ile 55	Ser	'Arg	Arg	Arg	Ġlu 60	His	Gly	Arg	Gly
20	Ser 65															
	(2)	INF	ORMA!	rion	FOR	SEQ	ID 1	NO: 1	L86:							
25			(i) :	(A) L B) T	ENGT YPE:	H: 2 ami	2 am no a	ino cid		s					
20			(xi)			OPOL				EQ I	D NO	: 18	6:			
30	Met	Val	Glu	Ser	Val	Met	Pro	Val	Val	Val	Cys	Thr	Leu	Ser	Pro	Gly
	1				5					10					15	
35	Ile	Asp	Ser	Ser 20	Pro	Ser		1				٠				ı
40	(2)	INF	ORMA:	SEQU (ENCE A) L	SEQ CHA ENGT YPE:	RACT H: 1	ERIS 32 a	TICS mino		ds					
45			(xi)			OPOL E DE				EQ I	D NO	: 18	7:			
	Met 1		Val	Leu	Phe 5	Val	Ala	Ile	Phe	Ala 10	Val	Pro	Leu	Ile	Leu 15	Gly
50	Gln	Glu	Tyr	Glu 20		Glu	Glu	Arg	Leu 25		Glu	Asp	Glu	Туг 30	Tyr	Gln
55	Val	Val	Tyr 35	Tyr	Tyr	Thr	Val	Thr 40	Pro	Ser	Tyr	Asp	Asp 45	Phe	Ser	Ala
	Asp	Phe 50	Thr	Ile	Asp	Tyr	Ser 55	Ile	Phe	G1u	Ser	Glu 60	Asp	Arg	Leu	Asn
60	Arg 65		Asp	Lys	Asp	Ile 70	Thr	Glu	Ala	Ile	Glu 75	Thr	Thr	Ile	Ser	Leu 80

	Glu Thr Ala Arg Ala Asp His Pro Lys Pro Val Thr Val Lys Pro Va 85 90 95
5	Thr Thr Glu Pro Gln Ser Pro Asp Leu Asn Asp Ala Val Ser Ser Let 100 105 110
10	Arg Ser Pro Ile Pro Leu Leu Ser Cys Ala Phe Val Gln Val Gly 115 120 125
	Met Tyr Phe Met 130
15	(2) INFORMATION FOR SEQ ID NO: 188:
20	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 69 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 188:
25	Met Pro Cys Gln Pro Gly Gln Val Pro Ser Cys Gln Cys Thr Phe Gly 1 5 10 15
	Leu Leu Leu Met Leu Pro Ser Leu Pro Ser Pro Ala Ser Gln Pro Arg 20 25 30
30	Pro Phe Cys Ser Ser Met Glu Tyr Phe His Gly Cys Ala Ser Pro Ser 35 40 45
35	Gln Ala Ile Ile Gly Gly Phe Pro Phe Ala Ser Val Ala Leu Ala Asp 50 55 60 Ile Leu Cys Leu Gln
40	(2) INFORMATION FOR SEQ ID NO: 189: (i) SEQUENCE CHARACTERISTICS:
45	(A) LENGTH: 45 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 189:
50	Met Ser Leu Leu Ser Pro Ala Ile Pro Ala Leu Thr Leu Ile Phe Ile 1 5 10 15
	Leu Met Phe Phe Ser Phe Pro Phe Arg Ala His Thr Val Val Thr Ile 20 25 30
55	Val Ala Ser Gly Phe Leu Gly Leu Ser Pro Leu Cys Gly 35 40 45
60	(2) INFORMATION FOR SEQ ID NO: 190:

			(i) :	C	A) L	ENGT	H: 6	5 am:	rics:		s			•		
5			(xi)	(1	D) T	OPOL	ami OGY: SCRI	line	ear	SQ II	D NO	: 190):			
10	Met 1	Ala	Phe	Gly	Leu 5	Gln	Met	Phe	Ile	Gln 10	Arg	Lys	Phe	Pro	Tyr 15	Pro
	Leu	Gln	Trp	Ser 20	Leu	Leu	Val	Ala	Val 25		Ala	Gly	Ser	Val 30	Val [.]	Ser
15	-		35			ı	J	40			1		45	Leu		
••	Phe	Leu 50	Glu	Thr	Gly	Gl'n	Leu 55	Pro	Lys	Asp	Arg	Ser 60	Thr	Asp	Gln	Arg
20	Ser 65					,				1					١	
25	(2)	INF					ID 1					i	•		1	
30				(A) I B) I D) I	ENGI YPE : YPOI	H: 5 ami OGY:	0 am no a lin	ear	acid): 19	1:			
35	Met 1	Asn	Leu	Leu	G1y 5		Ile	, Phe	Ser	Met 10	Cys	Ġļy	Leu	Met	Leu 15	Lys
	Leu	Lys	Trp	Cys 20		Trp	Val	Ala	Val 25	Tyr	Cys	Ser	Phe	Ile 30	Ser	Phe
40	Ala	Asn	Ser 35	-	Ser	Ser	Glu	Asp 40	Thr	Lys	Gln	Met	Met 45	Ser	Ser	Phe
45	Met	Хаа 50														
	(2)	INF	ORMA	TION	FOF	SEC) ID	NO:	192:							
50				_	(A) 1 (B) ' (D) '	LENG IYPE IOPO	TH: : : am: LOGY	170 a ino a : lir	near	ac:		D: 19	2:			
55	Met 1		ı Let	ı Asr	val		a Leu	ı Val	. Ala	Let 10		l Lev	Leu	Gly	Ala 15	_
60	Arg	Leu	ı Trp	Va] 20		Tr	Gly	/ Arg	Arg 25		, Leu	ı Gly	Ala	Gly 30		Gly

	Ala	Gly	Glu 35	Glu	Ser	Pro	Ala	Thr 40	Ser	Leu	Pro	Arg	Met 45	Lys	Lys	Arg
5	Asp	Phe 50	Ser	Leu	Glu	Gln	Leu 55	Arg	Gln	Туг	Asp	Gly 60	Ser	Arg	Asn	Pro
	Arg 65		. Leu	Leu	Ala	Val 70	Asn	Gly	Lys	Val	Phe 75	Ąsp	Vаl	Thr	Lys	Gly 80
10	Ser	Lys	Phe	Tyr	Gly 85	Pro	Ala	Gly	Pro	Tyr 90	Gly	Ile	Phe	Ala	Gly 95	Arg
15	Asp	Ala	Ser	Arg 100		Leu	Ala	Thr	Phe ,105	Cys	Leu	Asp 	Lys	Asp 110	Ala	Leu
10	Arg	Asp	Glu 115		Asp	Asp	Leu	Ser 120		Leu	Asn	Ala	Val 125	Gln	Met	Glu
20		130					135					140				i
	Gly 145		g Leu	. Leu	Lys	Pro 150		Glu	Glu	Pro	Ser 155		Tyr	Thr	Asp	Glu 160
25	Glu	ı Ası	o Thr	Lys	Asp 165		Asn	Lys	Gln	170						
30	(2)	IN	FORM	•												
35					(A) 1 (B) ' (D) '	LENG TYPE TOPO:	TH: (: am: LOGY	66 ar ino a : lir	mino acid near	aci		o: 1	93 : _.			
40	. :	1	т Ту		9	5				10)				15	5
			a Le	20)				25	5				30)	
45	Le	u Le	eu Gl		a Gly	y Cy:	s Ala	a Thi		e Lei	u Vai	l Thi	r Sei 4!		ı Ala	a Met
	Th		a As	p Le	u Il	e Gl	y Pro 5!		s Thi	r Ası	n. Se:	r Gly		u Sei	r Cy:	s Thi
50		a Pr 5	ro													
55	(2	:) II	NFORM	ATIO	n fo	r se	Q ID	NO:	194	:						
			(i)	SEC	(A)	LEN	TH:	92 a	mino	aci	ids					
60					(B) (D)			nino (: li								

30

(xi)	SEQUENCE	DESCRIPTION:	SEQ	ID	NO:	194:
------	----------	--------------	-----	----	-----	------

Met Ala Ala Gly Pro Ser Gly Cys Leu Val Pro Ala Phe Gly Leu Arg

1 5 10 15

Leu Leu Leu Ala Thr Val Leu Gln Ala Val Ser Ala Phe Gly Ala Glu
20 25 30

Phe Ser Ser Glu Ala Cys Arg Glu Leu Gly Phe Ser Ser Asn Leu Leu

35 40 45

Cys Ser Ser Cys Asp Leu Leu Gly Gln Phe Asn Leu Leu Gln Leu Asp 50 55 60

Pro Asp Cys Arg Gly Cys Cys Gln Glu Glu Ala Gln Phe Glu Thr Lys
65 70 75 80

Lys Leu Tyr Ala Gly Ala Ile Leu Glu Val Cys Gly 85 90

(2) INFORMATION FOR SEQ ID NO: 195:

25 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 176 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 195:

Met Arg Gly Ser His Leu Arg Leu Leu Pro Tyr Leu Val Ala Ala Asn

1 5 10 15

Pro Val Asn Tyr Gly Arg Pro Tyr Arg Leu Ser Cys Val Glu Ala Phe
20 25 30

Ala Ala Thr Phe Cys Ile Val Gly Phe Pro Asp Leu Ala Val Ile Leu
35 40 45

40 Leu Arg Lys Phe Lys Trp Gly Lys Gly Phe Leu Asp Leu Asn Arg Gln 50 55 60

Leu Leu Asp Lys Tyr Ala Ala Cys Gly Ser Pro Glu Glu Val Leu Gln
65 70 75 80

Ala Glu Glu Glu Phe Leu Ala Asn Ala Lys Glu Ser Pro Gln Glu Glu 85 90 95

Glu Ile Asp Pro Phe Asp Val Asp Ser Gly Arg Glu Phe Gly Asn Pro
100 105 110

Asn Arg Pro Val Ala Ser Thr Arg Leu Pro Ser Asp Thr Asp Asp Ser 115 120 125

Asp Ala Ser Glu Asp Pro Gly Pro Xaa Ala Glu Arg Gly Gly Ala Ser
130 135 140

Ser Ser Cys Cys Glu Glu Glu Gln Thr Gln Gly Arg Gly Ala Glu Ala 145 150 155 160

60

Arg Ala Pro Ala Glu Val Trp Lys Gly Ile Lys Lys Arg Gln Arg Asp

```
5
      (2) INFORMATION FOR SEQ ID NO: 196:
10
               . .
             (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 70 amino acids
                    (B) TYPE: amino acid
                   (D) TOPOLOGY: linear
15
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 196:
     Met Ser Asn Ala Cys Lys Glu Leu Ala Ile Phe Leu Thr Thr Gly Ile
                                          10
20
      Val Val Ser Ala Phe Gly Leu Pro Ile Val Phe Ala Arg Ala His Leu
                                      25 |
      Ile Glu Trp Gly Ala Cys Ala Leu Val Leu Thr Gly Asn Thr Val Ile
                                  40
25
      Phe Ala Thr Ile Leu Gly Phe Phe Leu Val Phe Gly Ser Asn Asp Asp
                                                  60
      Phe Ser Trp Gln Gln Trp
30
      (2) INFORMATION FOR SEQ ID NO: 197:
35
             (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 25 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
40
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 197:
      Met Thr Leu Leu Ile Ile Phe Leu Pro Phe Xaa Phe Thr Thr Xaa Thr
45
      Asn Ser Gly Gly Ser Phe Pro Val Arg
                  20
50
      (2) INFORMATION FOR SEQ ID NO: 198:
             (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 73 amino acids
                    (B) TYPE: amino acid
55
                    (D) TOPOLOGY: linear
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 198:
      Met Lys Gly Glu Leu Leu Pro Phe Leu Phe Leu Thr Val Trp Leu Trp
                                         10
                  5
60
```

	Leu	1 Туз	: Lys	Let 20		. Phe	e Gly	Glu	Ser 25		Arg	у Туз	Pro	Asn 30		. 11
5	Gly	/ Lys	Thr 35	Tyr	Phe	Phe	Phe	Trp 40		Asp	Glr	ıle	Ser 45		, Glu	Se
	Arg	Phe 50		Glu	Arg	Leu	Ala 55		Ile	Val	. Ser	Glu 60		Cys	Leu	11
10	Phe 65		ılle	His	Ala	Ile 70	Thr	Gly	Gln							
15	(2)	INF	ORMA	TION	FOR	SEQ	ID:	NO:	199:			i				
				SEQU	ENCE	СНА	RACT	ERIS	TICS		.ds					
20				((B) I (D) I	YPE: OPOI	ami OGY:	no a	cid ear							
							SCRI									
25	Met 1	Ser	Gly	Phe	Ser 5	Thr	Glu	Glu	Arg	Ala 10	Ala	Pro	Phe	Ser	Leu 15	Gl
	Tyr	Arg	Val	Phe 20	Leu	Lys	Asn	Glu	Lys 25	Gly	Gln	Tyr	Ile	Ser 30	Pro	Ph
30	His	Asp	Ile 35	Pro	Ile	Tyr	Ala	Asp 40	Lys	Asp	Val	Phe	His 45	Met	Val	Va.
	Glu	Val 50	Pro	Arg	Trp	Ser	Asn 55	Ala '	Lys	Met	Glu	Ile 60	Ala	Thr	Lys	Ası
35	Pro 65	Leu	Asn	Pro	Ile	Lys 70	Gln	Asp	Val	Lys	Lys 75	Gly	Lys	Leu	Arg	Ty:
40	Val	Ala	Asn	Leu	Phe 85	Pro	Тух	Lys	Gly	Tyr 90	Ile	Trp	Asn	Тут	Gly 95	Ala
	Ile	Pro	Gln	Thr 100	Trp	Glu	Asp	Pro	Gly 105	His	Asn	Asp	Lys	His 110	Thr	Gl
45	Cys	Cys	Gly 115	Asp	Asn	Asp	Pro	Ile 120	Asp	Val	Cys	Glu	Ile 125	Gly	Ser	Lys
	Val	Cys 130	Ala	Arg	Gly	Glu	Ile 135	Ile	Gly	Val	Lys	Val 140	Leu	Gly	Ile	Let
50	Ala 145	Met	Ile	Asp	Glu	Gly 150	Glu	Thr	Asp	Trp	Lys 155	Val	Ile	Ala	Ile	Asr 160
55	Val	Asp	Asp	Pro	Asp 165	Ala	Ala	Asn	Тут	Asn 170	Asp	Ile	Asn	Asp	Val 175	Lys
	Arg	Leu	Lys	Pro 180	Gly	Tyr	Leu	Glu	Ala 185	Thr	Val	Asp	Trp	Phe 190	Arg	Arg
60	Tyr	Lys	Val 195	Pro	Asp	Gly	Lys	Pro 200	Glu	Asn	Glu	Phe	Ala 205	Phe	Asn	Ala

	Glu	Phe 210	Lys	Asp	Lys	Asp	Phe 215	Ala	Ile	Asp	Ile	11e 220	Lys	Ser	Thr	His
5	Asp 225	His	Trp	Lys	Ala	Leu 230	Val	Thr	Lys	Lys	Thr 235	Asn	Gly	Lys	Gly	Ile 240
10	Ser	Cys	Met	Asn	Thr 245	Thr	Leu	Ser	Glu	Ser 250	Pro	Phe	Lys	Cys	Asp 255	Pro
10	Asp	Ala	Ala	Arg 260	Ala	Ile	Val	Asp	Ala 265	Leu	Pro	Pro	Pro	Cys 270	Ġlu	Ser
15	Ala	Cys	Thr 275	Val	Pro	Thr	Asp	Val 280	Ąsp	Lys	Trp	Phe	His 285	His	Gln	Lys
	Asn										•					
20																ı
	(2)	INF	ORMA!													
25			(1)	(ENCE A) L B) T D) T	ENGT YPE:	H: 6 ami	25 a no a	mino cid		ds					
				SEQ	UENC	E DE	SCRI	PTIO	N: S							
30	Met 1		Ile	Pro	Gly 5	Ser	Leu	Сув	Lys	Lys 10	Val	Lys	Leu	Ser	Asn 15	Asn
35	Ala	Gln	Asn	Trp 20	Gly	Met	Gln	Arg	Ala 25	Thr	Asn	Val	Thr	Tyr 30	Gln	Ala
33	His	His	Val 35	Ser	Arg	Asn	Lys	Arg 40	Gly	Gln	Val	Val	Gly 45	Thr	Arg	Gly
40	Gly	Phe 50		Gly	Cys	Thr	Val 55	Trp	Leu	Thr	Gly	Leu 60	Ser	Gly	Ala	Gly
	Lys 65		Thr	Val	Ser	Met 70	Ala	Leu	Glu	Glu	Tyr 75	Leu	Val	Cys	His	Gly 80
45	Ile	Pro	Cys	Tyr	Thr 85	Leu	Asp	Gly	Asp	Asn 90	Ile	Arg	Gln	Gly	Leu 95	Asn
50	Lys	Asn	Leu	Gly 100		Ser	Pro	Glu	Asp 105		Glu	Glu	Asn	Val 110	Arg	Arg
50	Ile	Ala	Glu 115		Ala	Lys	Leu	Phe 120		. Asp	Ala	Gly	Leu 125	Val	Cys	Ile
55	Thr	Ser 130		Ile	Ser	Pro	Туг 135		Gln	Asp	Arg	Asn 140		Ala	Arg	Gln
	Ile 145		Glu	Gly	Ala	Ser 150		Pro	Phe	Phe	Glu 155		Phe	Val	Asp	Ala 160
60	Pro	Lev	His	Val	. Cys	Glu	Gln	Arg	Asp	Val	Lys	Gly	Leu	Тут	Lys	Lys

					165					170		•	1		175	
· 5	Ala	Arg	Ala	Gly 180		Ile	Lys	Gly	Phe 185		Gly	Ile	Asp	Ser 190	Glu	Туз
	Glu	Lys	Pro 195	Glu	Ala	Pro	Glu	Leu 200		Leu	Lys	Thr	Asp 205		Cys	Asp
10	Val	Asn 210	Asp	Cys ,	Val	Gln	Gln 215	Val	Val	Glu	Leu '	Leu 220	Gln	Glu ,	Arg	Asp
	Ile 225		Pro	Val	Asp	Ala 230	Ser	Tyr	Glu	Val	Lys 235	Glu	Leu	Tyr	Val	Pro 240
15	Glu	Asn	Lys	Leu	His 245	Leu	Ala	Lys		Asp 250	Ala	Glu	Thr	Leu	Pro 255	Ala
20	Leu	Lys	Ile	Asņ 260	Lys	Val	Asp	Met	Gln 265	Trp	Val	Gln	Val	Leu 270	Ala	Glu
	Gly	Trp	Ala 275	Thr	Pro	Leu	Asn	Gly 280	Phe	Met	Arg	Glu	Arg 285	Glu	Tyr	Leu
25	Gln	Cys 290	Leu	His	Phe	Asp	Cys 295	Leu	Leu	Asp	Gly	Gly 300	Val	Ile	Asn	Leu
	Ser 305	Val	Pro	Įle	Val	Leu 310	Thr	Ala	Thr	His	Glu 315	Asp	Lys	Glu	Arg	Leu 320
30	Asp	Gly	Cys	Thr	Ala 325	Phe	Ala	Leu	Met	Туr 330	Glu	Gly	Arg	Arg	Val 335	Ala
35	Ile	Leu	Arg	Asn 340	Pro	Glu	Phe	.Phe	Glu 345	His	Arg	Lys	Glu	Glu 350	Arg	Суѕ
	Ala	Arg	Gln 355	Trp	Gly	Thr	Thr	Сув 360	Lys	Asn	His	Pro	Tyr 365	Ile	Lys	Met
40	Val	Met 370	Glu	Gln	Gly	Asp	Trp 375	Leu	Ile	Gly	Gly	Asp 380	Leu	Gln	Val	Leu
	Asp 385	Arg	Val	Tyr		Asn 390		Gly	Leu		Gln 395		Arg	Leu	Thr	Pro 400
45	Thr	Glu	Leu	Lys	Gln 405	Lys	Phe	Lys	Asp	Met 410	Asn	Ala	Asp	Ala	Val 415	Phe
50	Ala	Phe	Gln	Leu 420	Arg	Asn	Pro	Val	His 425	Asn	Gly	His	Ala	Leu 430	Leu	Met
	Gln	Asp	Thr 435	His	Lys	Gln	Leu	Leu 440	Glu	Arg	Gly	Tyr	Arg 445	Arg	Pro	Val
55	Leu	Leu 450	Leu	His	Pro	Leu	Gly 45 5	Gly	Trp	Thr	Lys	Asp 460	Asp	Asp	Val	Pro
	Leu 465	Met	Trp	Arg	Met	Lys 470	Gln	His	Ala	Ala	Val 475	Leu	Glu	Glu	Gly	Val 480
60	Leu	Asn	Pro	Glu	Thr	Thr	Val	Val	Ala	Ile	Phe	Pro	Ser	Pro	Met	Met

					485					490			t		495	
5	Tyr	Ala	Gly	Pro 500	Thr	Glu	Val	Gln	Trp 505	His	Cys	Arg	Ala	Arg 510	Met	Val
	Ala	Gly	Ala 515	Asn	Phe	Tyr	Ile	Val 520	Gly	Arg	Asp	Pro	Ala 525	Gly	Met	Pro
10	His	Pro 530	Glu	Ţhr '	Gly	Lys	Asp 535	Leu	Туг	Glu	Pro	Ser 540	His	Gly	Ala	Lys
	Val 545	Leu	Thr	Met	Ala	Pro 550	Gly	Leu	Ile	Ťhr	Leu 555	Glu	Ile	Val	Pro	Phe 560
15	Arg	Val	Ala	Ala	Туг 565	Asn	Lys	Lys		Lys 570	Arg	Met	Asp	Tyr	Tyr 575	Asp
20	Ser	Glu	His	Ніs 580	Glu	Asp	Phe	Glu	Phe 585	Ile	Ser	Gly	Thr	Arg 590	Met	Arg
	Lys	Leu	Ala 595	Arg	Glu	Gly	Gln	Lys 600	Pro	Pro	Glu	Gly	Phe 605	Met	Ala \	Pro
25	Lys	Ala 610	Trp	Thr	Val	Leu	Thr 615	Glu	Tyr	Tyr	Lys	Ser 620	Leu	Glu	Lys	Ala
20	Xaa 625															
30	(2)	******														
	(2)	INF	JRMA'	rion	FOR	SEQ	ID I	NO : 2	201:			(
35	(2)		(i)	SEQU)))	FOR ENCE A) L B) T D) T UENC	CHA ENGT YPE: OPOL	RACT H: 6 ami OGY:	ERIS 49 a no a lin	rics mino cid ear	aci		, 1	1:			
35 40			(i)	SEQU ((SEQ	ENCE A) L B) T D) T UENC	CHA ENGT YPE: OPOL E DE	RACT H: 6 ami OGY: SCRI	ERIS 49 a no a lin PTIO	TICS mino cid ear N: S	aci EQ I	D NO	: 20		Pro	Lys 15	Pro
	Met 1	Ser	(i) (xi)	SEQU ((SEQ Ser	ENCE A) L B) T D) T UENC Gln 5	CHA ENGT YPE: OPOL E DE	RACT H: 6 ami OGY: SCRI Leu	ERIS 49 a no a lin PTIO	rics mino cid ear N: S	aci EQ I Lys 10	D NO Pro	: 20 Leu	Phe		15	
40	Met 1 Ala	Ser Phe	(i) (xi) Ala	SEQU (((SEQ Ser Gln 20	ENCE A) L B) T D) T UENC Gln 5	CHA ENGT YPE: OPOL E DE Asp	RACT H: 6 ami OGY: SCRI Leu Pro	ERIS 49 a no a lin PTIO Glu Leu	rics mino cid ear N: S Pro	EQ I Lys 10 Thr	D NO Pro Glu	: 20 Leu Asn	Phe	His 30	15 Glu	Asp
40	Met 1 Ala Glu	Ser Phe Ser	(i) (xi) Ala Gly	SEQU ((SEQ Ser Gln 20	ENCE A) L B) T D) T UENC Gln 5 Lys	CHAI ENGT YPE: OPOL E DE Asp Pro	RACT H: 6 ami OGY: SCRI Leu Pro	ERIS 49 a no a lin PTIO Glu Leu Ser 40	rics mino cid ear N: S Pro Ser 25	EQ II Lys 10 Thr	D NO Pro Glu Lys	: 20 Leu Asn Gly	Phe Ser Ser 45	His 30 Pro	15 Glu Ala	Asp
40 45 50	Met 1 Ala Glu Leu	Ser Phe Ser Gly 50	(i) (xi) Ala Gly Pro 35	SEQU (() () () SEQ Ser Gln 20 Met	ENCE A) L B) T D) T UENC Gln 5 Lys Lys Ser	CHANGE ENGT YPE: OOPOL ASP Pro ASN	RACTH: 6 ami OGY: SCRI Leu Pro Val Ser 55	ERIS 49 a no a lin PTIO Glu Leu Ser 40	TICS mino cid ear N: S Pro Ser 25 Ser	EQ I Lys 10 Thr Ser	D NO Pro Glu Lys	: 20 Leu Asn Gly	Phe Ser Ser 45	His 30 Pro	15 Glu Ala Glu	Asp Pro
40 45	Met 1 Ala Glu Leu Ser 65	Ser Phe Ser Gly 50	(i) (xi) Ala Gly Pro 35	SEQU (() () () SEQ Ser Gln 20 Met Arg	ENCE A) L B) T D) T UENC Gln 5 Lys Lys Ser Asp	CHANGE ENGTYPE: OPPOL ASP Pro ASN Lys His 70	RACTH: 6 ami OGY: SCRI Leu Pro Val Ser 55 Ala	ERIS 49 a no a lin PTIO Glu Leu Ser 40 Gly	TICS mino cid ear N: S Pro Ser 25 Ser Pro Glu	EQ I Lys 10 Thr Ser	D NO Pro Glu Lys Lys Ser 75	: 20 Leu Asn Gly Pro 60 Ser	Phe Ser Ser 45 Ala	His 30 Pro Arg	15 Glu Ala Glu Phe	Asp Pro

	Ası	ı Thi	115	e Glr	ı Sei	: Lys	Ile	120		Glu	Glu	Leu	125		Gly	Thr
5	Pro	130	Ala	Arg	J Ph∈	Pro	Lys 135		Pro	Ser	Lys	Leu 140		Val	Gly	Gly
	Pro 145	Tr	Gly	Glr	Ser	150	Glu	Lys	Glu	Lys	Gly 155		Lys	Asn	Ser	Ala 160
10	Thr	Pro	Lys	Gln	165	Pro	Leu	Pro	Pro	Leu 170		Thr	Leu	Gly	Pro 175	Pro
15	Pro	Pro	Lys	180	Asn	Arg	Pro	Pro	Asn 185	Val	Asp	Leu	Thr	Lys 190	Phe	His
			195					Thr 200					205			
20		210					215	Pro				220				
05	225					230		Ser			235					240
25					245			Phe /		250					255	
30				260				His	265					270		
			275					Glu 280					285			
35		290					295					300				Glu ,
40	305					310		Glu			315					320
70					325			Gly		330					335	
45				340				Gly	345					350		
			355					Ile 360					365			
50		370					375	Arg				380				
55	385					390		Asp			395					400
JJ					405			Ile		410					415	
60	ASP	val	Ala	Glu 420	Gln	Asp .	Asp	Ile	Ser 4 25	Ser	His	Ser		Ser 430	Gly	Ser

	Gly	Gly	Ile 435	Phe	Pro	Pro	Pro	Pro 440	Asp	Asp	Asp	Ile	Tyr 445	Asp		Ile
5	Glu	Glu 450	Glu	Asp	Ala	Asp	Asp 455	Gly	Ser	Thr	Leu	ģln 460	Val	Gln	Glu	Lys
	Ser 465	Asn	Thr	Trp	Ser	Trp 470	-	Ile	Leu	Lys	475	Leu	Lys	Gly	Lys	Asp 480
10	Asp	Arg	Lys	Lys	Ser 485	Ile	Arg	Glu	Lys	Pro 490	Lys	Val	Ser	Asp	Ser 495	Asp
15	Asn	Asn	Glu	Gly 500	Ser	Ser	Phe	Pro	Ala 505	Pro	Pro	Lys	Gln	Leu 510	Asp	Met
13	Gly	Asp	Glu 515	Val	Tyr	Asp	Asp	Val 520	Asp'	Thr	Ser	Asp	Phe 525	Pro	Val	Ser
20	Ser	Ala 530	Glu	Met	Ser	Gln	Gly ,535	Thr	Asn	Val	Gly	Lys 540	Ala	Lys	Thr	Glu
	Glu 545	Lys	Asp	Leu	Lys	Lys 550	Leu	Lys	Lys	Gln	Хаа 555	Lys	Xaa	Xaa	Lys (Asp 560
25	Phe	Arg	Lys	Lys	Phe 565	Lys	Tyr	Asp	Gly	Glu 570	Ile	Arg	Val	Leu	Tyr 575	Ser
30	Thr	Lys	Val	Thr 580	Thr	Ser	Ile	Thr	Ser 585	Lys	Lys	Trp	Gly	Thr 590	Arg	Asp
	Leu	Gln	Val 595	Lys	Pro	Gly	Glu	Ser 600	Leu	Glu	Val	Ile '	Gln ,605	Thr	Thr	Asp
35	Asp	Thr 610	Lys	Val	Leu	Cys	Arg 615	Asn	Glu	Glu	Gly	Lys 620	Tyr	Gly	Tyr	Val
	Leu 625	Arg	Ser	Tyr	Leu	Ala 630	Asp	Asn	Asp	Gly	Glu 635	Ile	Tyr	Asp	Asp	Ile 640
40	Ala	Asp	Gly	Cys	Ile 645	Tyr	Asp	Asn	Asp							
45 ·	(2)	INF	ORMA	TION	FOR	SEQ	ID 1	NO:	202:							
50				- (((A) I (B) T (D) T	ENGI YPE : YPOL	H: 5 ami OGY:	ERIS 5 am no a lin	ino cid ear	ació						
	Mot	λla						PTIO						Pro	Val	T.en
55	1		_		5			-		10					15	
	cys	TÀL	rea	20		rea	пр	reu	25		FIIE	ser	пр	Ile 30	GIU	GIU
60	Leu	Lys	Ala 35		Leu	Arg	Asp	Asp 40		Leu	Ile	Ser	Ala 45	Val	Ala	Trp

Asn Ala Glu Phe Gln Thr Cys 50 55

5															,	
	(2)	INF	ORMA	TION	FOR	SEC	DI	NO:	203:				,			
10					(A) I (B) 1 (D) 1	ENG: TYPE: TOPOI	FH: 2 : ami	267 a ino a : lir	min cid ear	ac:): 20				
15	Met 1	Val	. Lys	Val	Thr 5		: Asn	Ser	Ala	Leu 10		Gln	Lys	Glu	Ala 15	_
20	Lys	Asp	Glu	Pro 20		Ser	Gly	Glu	Glu 25		Leu	Ile	Ile	Pro 30		As
	Ala	Val	Ala 35	Val	Asp	Cys	Lys	Asp 40	Pro	Asp	Asp	Val	Val 45		Val	Gl
25	Gln	Arg 50	Arg	Ala	Trp	Cys	Trp 55		Met	Cys	Phe	Gly 60		Ala	Phe	Me
	Leu 65	Ala	Gly	Val	Ile	Leu 70		Gly	Ala	Tyr	Leu 75		Lys	Tyr	Phe	A1 8
30	Leu	Gln	Pro	Asp	Asp 85	Val	Tyr	Tyr	Cys	Gly 90	Ile	Lys	Tyr	Ile	Lys 95	As
35	Asp	Val	Ile	Leu 100	Asn	Glu	Pro	Ser	Ala 105	Asp	Ala	Prọ	Ala	Ala 110		Тy
	Gln	Thr	Ile 115	Glu	Glu	Asn	Ile	Lys 120	Ile	Phe	Glu	Glu	Glu 125	Glu	Val	Gl
40	Phe	Ile 130	Ser	Val	Pro	Val	Pro 135	Glu	Phe	Ala	Asp	Ser 140	Asp	Pro	Ala	Ası
	Ile 145	Val	His	Asp	Phe	Asn 150	Lys	Lys	Leu	Thr	Ala 155	Tyr	Leu	Asp	Leu	As:
45	Leu	Asp	Lys	Cys	Туг 165	Val	Ile	Pro	Leu	Asn 170	Thr	Ser	Ile	Val	Met 175	Pro
50	Pro	Arg	Asn	Leu 180	Leu	Glu	Leu	Leu	Ile 185	Asn	Ile	Lys	Ala	Gly 190	Thr	Туз
	Leu	Pro	Gln 195	Ser	Tyr	Leu	Ile	His 200	Glu	His	Met	Val	Ile 205	Thr	Asp	Arg
55	Ile	Glu 210	Asn	Ile	Asp	His	Leu 215	Gly	Phe	Phe	Ile	Tyr 220	Arg	Leu	Cys	His
	Asp 225	Lys	Glu	Thr	Tyr	Lys 230	Leu	Gln	Arg	Arg	Glu 235	Thr	Ile	Lys	Gly	11e 240
60	Gln	Lys	Arg	Glu	Ala	Ser	Asn	Cys	Phe	Ala	Ile	Ara	His	Phe	Glu	Asr

				٠	245				•	250					255	
5	Lys	Phe	Ala	Val 260	Glu	Thr	Leu		Суs 265	Ser	Xaa				ı	
	(2)	INFC	RMAT	NOI	FOR	SEQ	ID N	ю: 2	04:	•		•	•			
10		ı	(i) S	(2	A) LI	ENGT	I: 3	ERIST 15 ar no ac	nino		ds					
			(xi)		•			line PTION		9Q II	NO:	204	ا ا:			
15	Met 1	Asp	Leu	Arg	Gln 5	Phe	Leu	Met	Cys	Leu 10	Ser	Leu	Cys	Thr	Ala 15	Phe
20	Ala	Leu	Ser	Lys 20	Pro	Thr	Glu	Lys	Lys 25	Asp	Arg	Val	His	His 30	Glu	Pro
	Gln	Leu	Ser 35	Asp	Lys	Val	His	Asn 40	Asp	Ala	Gln	Ser	Phe 45	Asp	Tyr	Asp
25	His	Asp 50	Ala	Phe	Leu	Gly	Ala 55	Gļu	Glu	Ala	Lys	Thr 60	Phe	Asp	Gln	Leu
30	Thr 65	Pro	Glu	Glu	Ser	Lys 70	Glu	Arg	Leu	Gly	Lys 75	Ile	Val	Ser	Lys	Ile 80
50	Asp	Gly	Asp	Lys	Asp 85	Gly	Phe	Val	Thr	Val 90	Asp	Glu	Leu	Lys	Asp 95	Trp
35	Ile	Lys	Phe	Ala 100	Gln	Lys	Arg	Trp	Ile 105	Tyr	Glu	Asp	Val	Glu 110	Arg	Gln
	Trp	Lys	Gly 115	His	Asp	Leu	Asn	Glu 120	Asp	Gly	Leu	Val	Ser 125	Trp	Glu	Glu
40	Туг	Lys 130		Ala	Thr	Tyr	Gly 135		Val	Leu	Asp	Asp 140	Pro	Asp	Pro	Asp
45	Asp 145	Gly	Phe	Asn	Tyr	Lys 150		Met	Met	Val	Arg 155	Asp	Glu	Arg	Arg	Phe 160
43	Lys	Met	Ala	Asp	Lys 165		Gly	' Asp	Leu	11e 170		Thr	Lys	Glu	Glu 175	Phe
50	Thr	Ala	Phe	Leu 180		Pro	Glu	Glu	Тут 185		Tyr	Met	Lys	Asp 190		Val
	Val	Gln	Glu 195		Met	Glu	Asp	1le 200		Lys	Asn	Ala	Asp 205		Phe	Ile
55	Asp	Lev 210		Glu	Tyr	Ile	Gly 215		Met	Туг	Ser	His 220		Gly	Asn	Thr
60	Asp 225		Pro	Glu	Trp	230		Thr	Glu	. Arg	Glu 235		Phe	Val	Glu	Phe 240

	Arg	j As	рL	ys	Asn	Arg 245		Gly	/ Lys	Met	250		Glu	Gļu	ı Thr	Lys 255	
5	Trj) Il	e L	eu	Pro 260		Asp	Туг	: Asp	His 265	Ala	Glu	Ala	ı' Glu	Ala 270	_	His
	Let	ı Va	1 T 2	у <u>г</u> 75	Glu	Ser	Asp	Gln	Asn 280		Asp	Gly	Lys	Leu 285		: Lys	Glu
10	Glu	11 29	e V O	al	Asp	Lys	Туг	Asp 295	Leu	Ph∈	· Val	Gly	Ser 300		Ala	, Thr	Asp
15	Phe 305		уG	lu	Ala	Leu	Val 310		His	Asp	Glu	Phe 315		í			
20	(2)	IN	(i) 5	SEQUI () ()	ENCE A) L B) T D) T	CHA ENGI YPE :	RACT H: 2 ami	no a	TICS mind cid	aci		•				
25	Met 1	Pho									EQ I Leu 10				Lys	Asp 15	Lys
30	Leu	Va:	l As	gz	Pro 20	Ile	Leu	Arg	Arg	His 25	Gly	Leu	Leu	Pro	Ser 30	Ser	Leu
	Lys	Arg	y II 3	le 35	Ala	Val	Gly	Met	Phe 40	Phe	Val	Met	Суş	Ser 45	Ala	Phe	Ala
35		50)					55			Asn		60				
40	65						70				Tyr	75					80
						85					Leu 90					9 5	
45				-	100					105	Ala				110		
50			11	.5					120		Phe	•		125			
50		130)					135			Ala		140				
55	Ile 145	Gly	Tr	i q	Met	Ser	Ser 150	His	Thr	Asp	Phe	Gly 155	Asn	Ile	Asn	Gly	Суз 160
	Tyr	Leu	. As	n I		Туг 165	Phe	Phe	Leu	Leu	Ala 170	Ala	Ile	Gln	Gly	Ala 175	Thr
60	Leu	Leu	Le		Phe : 180	Leu	Ile	Ile	Ser	Val 185	Lys	Tyr	Asp	His	His 190	Arg	Asp

	His	Gln	Arg 195	Ser	Arg	Ala	Asn	Gly 200	Val	Pro	Thr		Arg 205	Arg	Ala	
5																
	(2)	INFO	ORMAT	MOI	FOR	SEQ	ID N	ю: 2	206 :							
10			(i) :	ે (() ()	A) L B) T D) T	ENGT YPE : OPOL	H: 1: ami: OGY:	96 a no a lin	mino cid ear	aci		: 200	, 5:	ı		٠
15	Met 1	Arg	Ser	Arg	Ile 5	Arg	Glu	Phe	Asp	Ser '10	Ser	Thr	Leu	Asn	Glu 15	Ser
20	Val	Arg	Asn	Thr 20	Ίle	Met	Arg	Asp	Leu 25	Lys	Ala	Val	Gly	Lys 30	Lys	Ph€
20	Met	His	Val 35	Leu	Tyr	Pro	Arg	Lуs 40	Ser	Asn	Thr	Leu	Leu 45	Arg		Trr
25	Asp	Leu 50	Trp	Gly	Pro	Leu	Ile 55	Leu	Cys	Val	Thr	Leu 60		Leu	Met	Leu
	Gln 65	Arg	Asp	Ser	Ala	Asp 70	Ser	Glu	Lys	Asp	Gly 75	Gly	Pro	Gln	Phe	Ala 80
30	Glu	Val	Phe	Val	Ile 85	Val	Trp	Phe	Gly	Ala 90	Val	Thr	Ile	Thr	Leu 95	Ası
35	Ser	Lys	Leu	Leu 100	Gly	Gly	Asn	IÌe	Ser 105	Phe	Phe	Gln	'Ser	Leu 110	Cys	Va]
	Leu	Gly	Тут 115	_	Ile	Leu	Pro	Leu 120		Val	Ala	Met	Leu 125	Ile	Cys	Arg
40	Leu	Val 130	Leu	Leu	Ala	Asp	Pro 135		Pro	Val	Asn	Phe 140	Met	Val	Arg	Le
	Phe 145		Val	Ile	Val	Met 150		Ala	Trp	Ser	11e 155		Ala	Ser	Thr	Ala 16
45	Phe	Leu	Ala	. Asp	Ser 165		Pro	Pro	Asn	Arg 170		Ala	Leu	Ala	Val 175	Ту
50	Pro	Val	. Phe	Leu 180		Туг	Phe	Val	11e		Trp	Met	Ile	Leu 190		Ph
	Thr	Pro	Gln 195													
55	(2)	INF	ORMA													•
			(i)	_	(A) I	E CHA	rh: :	331 a	amino	o ac	ids					
60					(B) '	TYPE	: am	ino a	acid							

(D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 207:

5	Met 1		Lys	Asp	Gln 5		Val	Glu	Asn	Ile 10		Val	. Ser	Pro	Val 15	
	Val	Ala	Ser	Ser 20		Gly	Leu	Val	Ser 25		Gly	Gly	Lys	Ala 30	Thr	Thr
10	Ala	Ser	Gln 35	Ala	Lys	Ala	Val	Leu 40	Ser	Ala	Glu	Gln	Leu 45		Asp	Glu
15	Glu	Val 50	His	Ala	Gly	Leu	Gly 55		Leu	Leu	Arg	Ser 60		, Ser	Asn	Ser
	Thr 65	Ala	. Arg	Asn	Val	Thr 70		Lys	Leu	Gly	Ser 75	Arg	Leu	Tyr	Gly	Pro 80
20	Ser	Ser	Val	Ser	Phe 85	Ala	Asp	Asp	Phe	Val 90	Arg	Ser	Ser	Lys	Gln 95	His
	Tyr	Asn	Суз	Glu 100	His	Ser	Lys	Ile	Asn 105	Phe	Arg	Asp	Lys	Arg 110	Ser	Ala
25	Leu	Gln	Ser 115	Ile	Aşn	Glu	Trp	Ala 120	Ala	Gln	Thr	Thr	Asp 125	Gly	Lys	Leu
30	Pro	Glu 130	Val	Thr	Lys	Asp	Val 135	Glu	Arg	Thr	Asp	Gly 140	Ala	Leu	Leu	Val
	Asn 145	Ala	Met	Phe	Phe	Lys 150	Pro	His	Trp	Asp	Glu 155	Lys	Phe	His	His	Lys 160
35	Met	Val	Asp	Asn	Arg 165	Gly	Phe	Met	Val	Thr 170	Arg	Ser	Tyr	Thr	Val 175	Gly
	Val	Met	Met	Met 180	His	Arg	Thr	Gly	Leu 185	Tyr	Asn	Tyr	Tyr	Asp 190	Asp	Glu
40	Lys	Glu	Lys 195	Leu	Gln	Ile	Val	Glu 200	Met	Pro	Leu	Ala	His 205	Lys	Leu	Ser
45	Ser	Leu 210	Ile	Ile	Leu	Met	Pro 215	His	His	Val	Glu	Pro 220	Leu	Glu	Arg	Leu
	Glu 225	Lys	Leu	Leu	Thr	Lys 230	Glu	Gln	Leu	Lys	Ile 235	Trp	Met	Gly	Lys	Met 240
50	Gln	Lys	Lys	Ala	Val 245	Ala	Ile	Ser	Leu	Pro 250	Lys	Gly	Val	Val	Glu 255	Val
	Thr	His	Asp	Leu 260	Gln	Lys	His	Leu	Ala 265	Gly	Leu	Gly	Leu	Thr 270	Glu	Ala
55	Ile	Asp	Lys 275	Asn	Lys	Ala	Asp	Leu 280	Ser	Arg	Met	Ser	Gly 285	Lys	Lys	Asp
60	Leu	Туr 290	Leu	Ala	Ser	Val	Phe 295	His	Ala	Thr	Ala	Phe 300	Glu	Leu	Asp	Thr

	305	027			200	310	9			GIY	315	Gly	Vai	ALG	1111	320
5	Val	Phe	Tyr	Ala	Asp 325	His	Pro	Phe	Ile	Ser 330	Xaa				•	
										'			. '			
10	(2)	INF						NO: 3			•					
			(i)	. (A) L B) T	ENGT YPE:	H: 5 ami	ERIS 8 am no a lin	ino cid		s	ı		ı	,	
15			(xi)					PTIO		BQ I	D NO	: .20	8:			
	Met 1	Cys	Met	Gln	Leu 5	Phe	Gly	Phe	Leu	Ala 10	Phe	Met	Ile	Phe	Met 15	Cys
20	Trp	Val	Gly	Asp 20	Val	Tyr	Pro	Val	Тут 25	Gln	Pro	Val	Gly	Pro 30	Lys	Gln
25	Tyr	Pro	Tyr 35	Asn	Asn	Leu	Tyr	Leu 40	Glu	Arg	Gly	Gly	Asp 45	Pro	Ser	Lys
	Glu	Pro 50	Glu	Arg	Val	Val	His 55	Tyr	Glu	Ile						
30	(2)	INFO	ORMA!	rion	FOR	SEQ	ID I	NO: 2	209:							
35			(i) :	(A) L	ENGT	H: 3	ERIS 92 a no a	mino		ds					
<i></i>			(xi)	(D) T	OPOL	OGY:	lin PTIO	ear	EQ I	D NO	: 20	9:			
40	Met 1	Asp	Ala	Leu	Val 5	Glu	Asp	Asp	Ile	Cys 10	Ile	Leu	Asn	His	Glu 15	Lys
	Ala	His	Lys	Arg 20	Asp	Thr	Val	Thr	Pro 25	Val	Ser	Ile	Tyr	Ser 30	Gly	Asp
45	Glu	Ser	Val 35	Ala	Ser	His	Phe	Ala 40	Leu	Val	Thr	Ala	Туг 45	Glu	Asp	Ile
50	Lys	Lys 50	Arg	Leu	Lys	Asp	Ser 55	Glu	Lys	Glu	Asn	Ser 60	Leu	Leu	Lys	Lys
J O	Arg 65	Ile	Arg	Phe	Leu	Glu 70	Glu	Lys	Leu	Ile	Ala 75	Arg	Phe	Glu	Glu	Glu 80
55	Thr	Ser	Ser	Val	Gly 85	Arg	Glu	Gln	Val	Asn 90	Lys	Ala	Tyr	His	Ala 95	Tyr
	Arg	Glu	Val	Cys 100	Ile	Asp _.	Arg	Asp	Asn 105	Leu	Lys	Ser	Lys	Leu 110	Asp	Lys
60	Met	Asn	Lys	Asp	Asn	Ser	Glu	Ser	Leu	Lys	Val	Leu	Asn	Glu	Gln	Leu

PAICHOCID: JAIO 00/2720/1 1 -

			115					120					12,5			
5	Gln	Ser 130	Lys	Glu	Val	Glu	Leu 135	Leu	Gln	Leu	Arg	Thr 140		Val	Glu	Thr
J	Gln 145	Gln	Val	Met	Arg	Asn 150	Leu	Asn	Pro	Pro	Ser 155	Ser	Asn	Trp	Glu	Val 160
10	Glu	Lys	Leu	Ser	Cys 165	Asp	Lėu	Lys	Ile	His 170	Gly	Leu	Glu	Gln	Glu 175	Leu
	Glu	Leu	Met	Arg 180	Lys	Glu	Cys	Ser	Asp 185	Leu	Lys	Ile	Glu	Leu 190	Gln	Lys
15	Ala	Lys	Gln 195		Asp	Pro	Tyr	Gln 200	Glu	Asp '	Asn		Lys 205	Ser	Arg	Asp
20	Leu	Gln 210	Lys	Leu	Şer	Ile	Ser 215	Ser	Asp	Asn	Met	Gln 220	His	Ala	Tyr	Trp
	Glu 225	Leu	Lys	Arg	Glu	Met 230	Ser	Asn	Leu	His	Leu 235	Val	Thr	Gln	Val	Gln 240
25	Ala	Glu	Leu	Leu	Arg 245	Lys	Leu	Lys	Thr	Ser 250	Thr	Ala	Ile	Lys	Lys 255	Åla
	Cys	Ala	Pro	Val 260	Gly	Cys	Ser	Glu	Asp 265	Leu	Gly	Arg	Asp	Ser 270	Thr	Lys
30	Leu	His	Leu 275	Met	Asn	Phe	Thr	Ala 280	Thr	Tyr	Thr	Arg	His 285	Pro	Pro	Leu
-35	Leu	Pro 290	Asn	Gly	Lys	Ala	Leu 295	Cys	His	Thr	Thr	Ser 300	Ser	Pro	Leu	Pro
	Gly 305	Asp	Val	Lys	Val	Leu 310	Ser	Glu	Lys	Ala	Ile 315	Leu	Gln	Ser	Trp	Thr 320
40	Asp	Asn	Glu	Arg	Ser 325	Ile	Pro	Asn	Asp	Gly 330	Thr	Cys	Phe	Gln	Glu 335	His
	Ser	Ser	_	Gly 340	_		Ser			_	Asn	Ser	Trp	Val 350	Phe	Pro
45	Ser	Pro	Pro 355		Ser	Ser	Glu	Thr 360	Ala	Phe	Gly	Glu	Thr 365	Lys	Thr	Lys
50	Thr	Leu 370		Leu	Pro	Asn	Leu 375	Pro	Pro	Leu	His	Tyr 380	Leu	Asp	Gln	His
	Asn 385	Gln	Asn	Cys	Leu	Tyr 390	Lys	Asn								
55	(2)	INF	ORMA	TION	FOR	SEQ	ID :	NO:	210:							
			(i)	SEQU												
60					(A) I (B) 7					ació	ls					

```
(D) TOPOLOGY: linear
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 210:
     Met His His His Thr Gln Leu Met Phe Ile Tyr Leu Phe Ile Tyr Leu
5
     Phe Ile Leu Gly Val Phe Phe Phe Phe Xaa
10
      (2) INFORMATION FOR SEQ ID NO: 211:
             (i) SEQUENCE CHARACTERISTICS:
15
                    (A) LENGTH: 39 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 211:
20
     Met Asn Cys Ile Leu Leu Leu Tyr Leu Leu Ile Pro Thr Ile Ser Ile
      Ser Val Val Pro Tyr Val Ala Leu Asn Ile Lys Tyr Ile Lys Glu Cys
                                      25
25
      Thr Glu Asn Ser Phe Tyr Xaa
               35
30
      (2) INFORMATION FOR SEQ ID NO: 212:
             (i) SEQUENCE CHARACTERISTICS:
                     (A) LENGTH: 71 amino acids
35
                     (B) TYPE: amino acid
                     (D) TOPOLOGY: linear
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 212:
      Met Leu Leu His Leu Thr Ala Ala Phe Leu Gln Arg Ala Gln Phe Ser
40
      Thr Tyr Phe Pro Gly Tyr Phe Asp Gly Gln Tyr Trp Leu Trp Trp Val
                                       25
45
      Phe Leu Val Leu Gly Phe Leu Leu Phe Leu Arg Gly Phe Ile Asn Tyr
      Ala Lys Val Arg Lys Met Pro Glu Thr Phe Ser Asn Leu Pro Arg Thr
50
      Arg Val Leu Phe Ile Tyr Xaa
       65
55
      (2) INFORMATION FOR SEQ ID NO: 213:
              (i) SEQUENCE CHARACTERISTICS:
                     (A) LENGTH: 83 amino acids
60
                     (B) TYPE: amino acid
```

	(D) TOP	OLOGY:	linear	r			'
i)	SEQUENCE 1	DESCRIP	TION:	SEQ	ID	NO:	213

Met Leu Thr Phe Phe Met Ala Phe Leu Phe Asn Trp Ile Gly Phe Phe 5 10

Leu Ser Phe Cys Leu Thr Thr Ser Ala Ala Gly Arg Tyr Gly Ala Ile 25

10 Ser Gly Phe Gly Leu Ser Leu Ile Lys Trp Ile Leu Ile Val Arg Phe 40

Ser Thr Tyr Phe Pro Ala Phe Met Asn Ser Leu Ser Arg Ser Lys Arg

Thr Pro Ala Gly Ser Glu Ser Arg Cys Arg Thr Gln Arg Asn Asn His 70 ' 75

Leu Leu Xaa

20

25

15

(2) INFORMATION FOR SEQ ID NO: 214:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 81 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear
- 30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 214:

Met Ser Lys Arg Ser Ala Ser Phe Ile Leu Leu Pro Leu Leu Phe Leu 1 5 10

35 Lys Gly Ser Phe Ala Lys Leu Asn Ala Arg Ile Ser Asp Cys Leu Glu

Glu Arg Tyr Cys His Asn Leu Trp Met Val Phe Gln Gly Cys Val Ile 35 40

Thr Glu Leu His Leu Ser Arg Met Ser Lys Thr Leu Ser Ser Leu Cys

Tyr Asp Phe Val Ile Asn Val Tyr Ile Phe Phe Lys Phe Leu Asp Ile 45

Thr

50

- (2) INFORMATION FOR SEQ ID NO: 215:
- (i) SEQUENCE CHARACTERISTICS: 55
 - (A) LENGTH: 49 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 215:
- 60 Met Cys Ser Leu Phe Glu Ser Arg Phe Phe Cys Phe Val Leu Phe Ser

WO 98/42738 PCT/US98/05311

	1				5					10			•		15	
5	Glu	Lys	Ile	Ile 20	Gln	Leu	Cys	Ala	Ser 25	Ile	Ala	Pḥe	Leu	Cys 30	Phe	Val
3	Lys	His	Val 35	Pro	Trp	Pro	Lys	Trp 40	Lys	Arg	Lys	Суѕ	Leu 45	Ile	Asn	Ala
10	Phe				,		,				ı		ı	•		
15	(2)	INFO	PAMAC	LION	FOR	SEQ	'ID I	NO: 2	216:	٠	,		•			
			(i) :	(. (:	A) Li b) T	ENGT YPE:	H: 2 ami	03 a no a			ds					
20			(xi)					lin PTIO	N: S	EQ I	D NO	: 21	6:			
	Met 1	Thr	Leu	Arg	Pro 5	Ser	Leu	Leu	Pro	Leu 10	His	Leu	Leu	Leu	Leu 15	Leu
25	Leu	Leu	Ser	Ala 20	Ala	Val	Cys	Arg	Ala 25	Ģlu	Ala	Gly	Leu	Glu 30	Thr	Glu
20	Ser	Pro	Val 35	Arg	Thr	Leu	Gln	Val 40	Glu	Thr	Leu	Val	Glu 45	Pro	Pro	Glu
30	Pro	Cys 50		Glu	Pro	Ala	Ala 55	Phe ,	Gly	Asp	Thr	Leu 60		Ile	His	Tyr
35	Thr 65	Gly	Ser	Leu	Val	Asp 70	Gly	Arg	Ile	Ile	Asp 75	Thr	Ser	Leu	Thr	Arg 80
	Asp	Pro	Leu	Val	Ile 85	Glu	Leu	Gly	Gln	Lys 90		Val	Ile	Pro	Gly 95	Leu
40	Glu	Gln	Ser	Leu 100	Leu	Asp	Met	Суз	Val 105	Gly	Glu	Lys	Arg	Arg 110		Ile
45	Ile	Pro	Ser 115		Leu	Ala	Туг	Gly 120		Arg	Gly	Phe	Pro 125	Pro	Ser	Val
43	Pro	Ala 130		Ala	Val	Val	Gln 135		Asp	Val	. Glu	Leu 140		Ala	Leu	Ile
50	Arg 145		Asn	Тут	Trp	Leu 150		Leu	Val	Lys	Gly 155		e Lev	Pro	Leu	Val 160
	Gly	Met	: Ala	Met	Val 165		Pro	Ser	Trp	Ala 170		Leu	ı Gly	/ Ile	Thr 175	Tyr
55	Thr	Glu	a Arg	180		Asp	Pro	Lys	Ser 185		Lys	Arg	g Ser	Ser 190		Lys
60	Arg) Asr	195		Arg	Ala	Lys	200	j Asn	a Ası	ı Lys	3				

	(2)) IN	FORM	TION	FOR	SEC) ID	NO:	217:							
5			(i)		(A) I (B) 1	LENG:	TH: :	186 a ino a	mino acid		ids				ì	
			(xi)	SEC	(D) 1 OUENC					ו ספו	TD NY)· 21				
10	Met	: Lys		Leu		Thr					Thr			ı Leu	ا Vạl 15	
15	Ser	: Ile	e Ser	Leu 20	Trp	Ile	Ile	Ala	Ala 25		Thr	Val	Arg	Val		G1
	Ser	Pro	35	Ser	Pro	Ala	Gln	Pro 40	Ser	Gly	Ser	Ser	Leu 45		Ala	Tr
20	Тух	His 50	Asp	Gln	Gln	Asp	Val 55	Thr	Ser	Asn	Phe	Leu 60	Gly	Ala	Met	Tr
25	Leu 65	Ile	: Ser	Ile	Thr	Phe 70	Leu	Ser	Ile	Gly	Туг 75	Gly	Asp	Met	Val	Pr
	His	Thr	Tyr	Cys	Gly 85	Lys	Gly	Vaļ	Cys	Leu 90	Leu	Thr	Gly	Ile	Met 95	G1
30	Ala	Gly	Cys	Thr 100	Ala	Leu	Val	Val	Ala 105	Val	Val	Ala	Arg	Lys 110	Leu	Glı
	Leu	Thr	Lys 115	Ala	Glu	Lys	His	Val 120	His	Xaa	Phe	Met	Met 125	Asp	Thr	Glr
35	Leu	Thr 130	Lys	Arg	Ile	Lys	Asn 135	Xaa	Ala	Ala	Asn	Val 140	Leu	Xaa	Glu	Thi
40	Trp 145	Leu	Ile	Tyr	Lys	His 150	Thr	Lys	Leu	Leu	Lys 155	Lys	Ile	Asp	His	Ala 160
	Lys	Val	Arg	Asn	Thr 165	Arg	Gly	Ser	Ser	Ser 170	Lys	Tyr	Pro	Pro	Val 175	Glu
45	Glu	Arg	Gln	Asp 180	Gly	Thr	Glu	Glu	Ala 185	Glu						
50	(2)			'ION SEQUE	NCE	CHAF	ACTE	RIST	ICS:							
					1) LE 3) TY					cids	3					
55		,	(xi)) TC	POLO	GY:	line	ar	QIC	NO:	218	:			
	Met 1	Lys	Phe	Leu .	Ala ' 5	Val:	Leu '	Val :	Leu :	Leu 10	Gly	Val	Ser	Ile	Phe 15	Leu
60	Val	Ser	Ala	Gln i	Asn 1	Pro '	Thr '	Thr :	11a :	λla	Dro	- ומ	3	mb	m	-

				20	•				25	1				30		
5	Ala	Thr	Gly 35	Pro	Ala	Asp	Asp	Glu 40	Ala	Pro	Asp	Ala	Glu 45	Thr	Thr	Ala
3	Ala	Ala 50	Thr	Thr	Ala	Thr	Thr 55	Ala	Ala	Pro	Thr	Thr 60	Ala	Thr	Thr	Ala
10	Ala 65	Ser	Thr	Thr	Ala	Arg 70	Lys	Asp	Ile	Pro	Val 75	Leu	Pro	Lys	Trp	Val 80
	Gly	Asp	Leu	Pro.	Asn 85	Gly	Arg	Val	Cys	Pro 90					٠	
15	(2)	INF	ORMA'	rion	FOR	SEQ	ID I	No: 2	219:							
20				(A) L B) T D) T	ENGT YPE: OPOL	H: 1 ami OGY:	39 a no a lin		aci		: 21	9:			
25	Met 1	Ser	Ser	Ala	Ala 5	Ala	Asp	His	Trp	Ala 10	Trp	Leu	Leu	Val	Leu 15	Ser
30	Phe	Val	Phe	Gly 20	Cys	Asn	Val	Leu	Arg 25	Ile	Leu	Leu	Pro	Ser 30	Phe	Ser
	Ser	Phe	Met 35	Ser	Arg	Val	Leu	Gln 40	Lys	Asp	Ala	Glu	Gln 45	Glu	Ser	Gln
35	Met	Arg 50	Ala	Glu	Ile	Gln	Asp 55	Met	Lys	Gln	Glu	Leu 60	Ser	Thr	Val	Asn
	Met 65	Met	Asp	Glu	Phe	Ala 70	Arg	Tyr	Ala	Arg	Leu 75	Glu	Arg	Lys	Ile	Asn 80
40					85				His	90					95	
45				100					Ser 105					110		
			115					120	Trp			_	Ser 125		Pro	Val
50	Ala	Val 130		Pro	Ser	Lys	Trp 135		Thr	Leu	Xaa					
55	(2)	INF		SEQU (ENCE (A) I (B) I		RACI TH: 4	ERIS 18 an ino a	TICS nino acid		ls					
60			(vi)	SEC	i ienc	म्या ज	SCRI	PTTC	N: 5	EO T	א ס	1. 22	n -			

	Met 1	Ser	Ser	Ala	Ala 5	Ala	Asp	His	Trp	Ala 10	Trp	Leu	Leu	Val	Leu 15	Ser
5	Phe	Val	Phe	Gly 20	Cys	Asn	Val	Leu	Arg 25	Ile	Leu	Leu	Pro	Ser 30	Phe	Ser
10	Ser	Phe	Met 35	Ser	Arg	Val	Leu	Gln 40	Lys	Asp	Ala	Asp	Arg 45	Ser	His	Arg
15	(2)	TME	О₽Ма	PION	EOP.	CTF.O.	, The s	NO - 1	221.							
20	(2)	TIME	(i)	SEQU (ENCE A) L B) T D) T	CHA ENGT YPE: OPOL	RACT H: 7 ami OGY:	ERIS O am no a lin	TICS ino cid ear	: acid		: 22	1:			
25	Met 1	Thr	Ala	Pro	Leu 5	Pro	Pro	Leu	Ser	Gly 10	Leu	Ala	Leu	Phe	Leu 15	Ile
	Val	Phe	Phe	Ser 20	Leu	Gly	Val	Phe	Cys 25	Ile	Cys	His	Ser	His '30	Trp	Tyr
30	His	Thr	Leu 35	Gln	Gln	Met	Ala	Gly 40	Thr	Glu	Pro	Lys	Ala 45	Leu	Leu	Leu
35		50					Thr 55	Phe	Val	Thr	Val	Thr 60	His	Glu	Val	Trp
	65 65	GIU	GIN	Ala	Leu	70										
40	(2)	INFO	ORMA?	rion	FOR	SEQ	ID 1	10: 2	222:							
45				C	A) L: B) T D) T	ENGT YPE : OPOL	H: 8: ami: OGY:	3 am no a lin	ino a cid ear	acid		: 222	2:			
50	Met 1	Thr	Суз	Ser	Val 5	Ala	Leu	Leu	Leu	Ile 10	Leu	Gly	Leu	Arg	Суз 15	Ser
	Gly	Val	Arg	Pro 20	Gly	Leu	Val	Gly	Glu 25	Gly	His	Asn	Pro	Ser 30	Leu	Leu
55	Val	Cys	Leu 35	Leu	Leu	Lys	Asp	Ser 40	Arg	Thr	Asn	Gln	Gly 45	Ser	Cys	Pro
50	Gly	Gly 50	Pro	Trp	Ser	Glu	Arg 55	Asp	Ile	Glu	Ser	Val 60	Thr	Ser	Asp	Asn

	Cys 65	Glu	Ala	Thr	Leu	Gly 70	Tyr	Arg	Asn	His	Ser 75	Leu	Pro	Ser	Asn	Tyr 80
5	туг	Asn	Ser									. '				
10	(2)			SEQUI	FOR ENCE A) L B) T	CHAI ENGT YPE:	RACT H: 4	ERIS 3 am no a	rics ino cid		ı s		,	•		
15			(xi)	-	D) T UENC					₽Q II	D NO	: 22	3:			
	Met 1	Leu	Thr	Arg	Ser 5	Leu	Lys	Thr	Leu	Pro 10	Ser	Ala	Cys	Thr	Ala 15	Phe
20	Leu	Leu	Leu	Phe 20	Phe	Leu	Phe	Ser	Ser 25	Gly '	Asp	Pro	Glu	Leu 30	Ser	Cys
25	Ser	Cys	Thr 35	Leu	Arg	Thr	Gln	Ser 40	Ser	Trp	Ser				,	
	(2)	INF	ORMA!	rion	FOR	SEQ	ID I	1 0: 2	224:					,		
30			(i)	SEQU	ENCE	CHA	RACT	ERIS	rics	:						
			(xi)	(A) L B) T D) T	YPE: OPOL	ami OGY:	no a lin	cid ear			. 22	, ' 4 •			
35	Met 1			SEQ	B) T D) T UENC	YPE: OPOL E DE:	ami OGY: SCRI	no a lin PTIO	cid ear N: S	EQ I	D NO	: 22		Arg	His 15	Gly
	1	Trp	Arg	(SEQ Pro	B) T D) T UENC Ser 5	YPE: OPOL E DE: Val	ami OGY: SCRI Leu	no a lin PTIO Leu	cid ear N: S Leu	EQ II Leu 10	D NO Leu	: 22 Leu	Leu		15	Gly Arg
35	1 Ala	Trp Gln	Arg Gly	((SEQ) Pro Lys 20	B) T D) T UENC Ser 5	YPE: OPOL E DE: Val Ser	ami OGY: SCRI Leu Pro	no a lin PTIO Leu Asp	cid ear N: S Leu Ala 25	EQ II Leu 10 Gly	D NO Leu Pro	: 22 Leu His	Leu Gly	Gln 30	15 Gly	Arg
35	1 Ala Val	Trp Gln His	Arg Gly Gln 35	() SEQUENTS Pro Lys 20 Ala	B) T D) T UENC Ser 5 Pro	YPE: OPOL E DE: Val Ser	ami OGY: SCRI Leu Pro	no a lin PTIO Leu Asp Ser 40	cid ear N: S Leu Ala 25	EQ II Leu 10 Gly Ala	D NO Leu Pro	: 22 Leu His	Leu Gly Asp 45	Gln 30 Asp	15 Gly Ala	Arg
35 40 45	1 Ala Val Gly	Trp Gln His Asn 50	Arg Gly Gln 35 Phe	(((SEQ) Pro Lys 20 Ala	B) T D) T UENC: Ser 5 Pro Ala	YPE: OPOL E DE: Val Ser Pro	ami OGY: SCRI Leu Pro Leu His 55	no a lin PTIO Leu Asp Ser 40	cid ear N: S Leu Ala 25 Asp	Leu 10 Gly Ala	D NO Leu Pro Pro	: 22 Leu His His Gly 60	Leu Gly Asp 45 Arg	Gln 30 Asp Glu	15 Gly Ala Val	Arg His
35 40	l Ala Val Gly Lys 65	Trp Gln His Asn 50 Glu	Gly Gln 35 Phe	(((SEQ) Pro Lys 20 Ala Gln	B) T D) T UENC: Ser 5 Pro Ala	YPE: OPOL E DE: Val Ser Pro Asp Leu 70	ami OGY: SCRI Leu Pro Leu His 55	no a lin PTIC Leu Asp Ser 40 Glu	cid ear N: S Leu Ala 25 Asp Ala Glu	Leu 10 Gly Ala Phe	D NO Leu Pro Pro Leu Ser 75	: 22 Leu His His Gly 60	Leu Gly Asp 45 Arg	Gln 30 Asp Glu Arg	15 Gly Ala Val Leu	Arg His Ala Gly 80
35 40 45	Ala Val Gly Lys 65 Arg	Trp Gln His Asn 50 Glu	Arg Gly Gln 35 Phe Val	(((SEQ) Pro Lys 20 Ala Gln Asp	B) T D) T UENC: Ser 5 Pro Ala Tyr Gln Arg	YPE: OPOL E DE: Val Ser Pro Asp Leu 70	ami OGY: SCRI Leu Pro Leu His 55 Thr	no a lin PTIO Leu Asp Ser 40 Glu Pro	cid ear N: S Leu Ala 25 Asp Ala Glu Ala	Leu 10 Gly Ala Phe Glu Gly 90	D NO Leu Pro Pro Leu Ser 75	E 22 Leu His His Gly 60 Gln Gly	Leu Gly Asp 45 Arg Ala	Gln 30 Asp Glu Arg	15 Gly Ala Val Leu Trp 95	Arg His Ala Gly 80 Val
35 40 45 50	Ala Val Gly Lys 65 Arg Ser	Trp Gln His Asn 50 Glu Ile	Arg Gly Gln 35 Phe Val	(((SEQQ Pro 20 Ala Gln Asp Glu 100	B) TD) TO UENCE Ser 5 Pro Ala Tyr Gln Arg 85	YPE: OPOL E DE: Val Ser Pro Asp Leu 70 Met	ami OGY: SCRI Leu Pro Leu His 55 Thr Asp	no a lin PTIO	cid ear N: S Leu Ala 25 Asp Ala Glu Ala Ile 105	Leu 10 Gly Ala Phe Glu Gly 90 Ala	D NO Leu Pro Pro Leu Ser 75 Asp	His Gly 60 Gln Gly	Leu Gly Asp 45 Arg Ala Asp	Gln 30 Asp Glu Arg Gly	15 Gly Ala Val Leu Trp 95 Arg	Arg His Ala Gly 80 Val

		130			•		135			•		140				
5	Xaa 145	Xaa	Pro	Xaa	Glu	Glu 150	Phe	His	Asp	Val	Glu 155	Asp	Ala	Glu	Thr	Туг 160
3	Lys	Lys	Met	Leu	Хаа 165	Arg	Asp	Glu	Arg	Arg 170	Phe	Arg	Val	Ala	Asp 175	Gln
10	Asp	Gly	Asp	Ser 180	Met	Ala	Thr	Arg								
					·										•	
15	(2)		ORMA			_			ı			ı	·			
			(1)	(. (:	A) L B) T	ENGT YPE:	H: 7 ami	ERIS 1 am no a lin	ino cid		s					
20			(xi)		-			PTIO		EQ I	D NO	: 22	5:			
	Met 1	Trp	Leu	Phe	Ile 5	Leu	Leu	Ser	Leu	Ala 10	Leu	Ile	Ser	Asp	Ala 15	Met
25	Val	Met	Asp	Glu 20	Lys	Val	Lys	Arg /	Ser 25	Leu	Cys	Trp	Thr	Arg 30	Leu	Leu
30	Pro	Ser	Ala 35	Thr	Thr	Met	Pro	Xaa 40	Thr	Arg	Ile	Thr	Pro 45	Asn	Thr	Gly
	Ala	Glu 50	Xaa	Ile	Ser	Val	Хаа 55	Thr	Ala	Thr	Ser	Ser 60	Pro	Ser	Pro	Leu
35	Thr 65	Ala	Pro	Ile	Met	Trp 70	Pro	ŀ					•			
40	(2)	INF	ORMA:	rion	FOR	SEQ	ID I	NO: 2	226:							
			(i) :	(.	A) L	engt	н: 1	ERIS 0 am no a	ino		s					
45			(xi)					lin PTIO		EQ I	D NO	: 22	6:			
	Met 1	His	Val	Phe	Val 5	Leu	Glu	Ile	Phe	Leu 10						
50																
	(2)	INF	ORMA!	rion	FOR	SEQ	ID I	NO: 2	227:					•		
55			(i)	(.	A) L	ENGT	H: 1	ERIS 38 a no a	mino		ds					
			(xi)	(D) T	OPOL	OGY:	no a lin PTIO	ear	EQ I	D NO	: 22	7:			
60	Met	Ala	Val	Ala	Thr	Leu	Ala	Ser	Glu	Thr	Leu	Pro	Leu	Leu	Ala	Leu

	1		•		5					10			1		15	
5	Thr	Phe	Ile	Thr 20	Asp	Asn	Ser	Leu	Val 25	Ala	Ala	Gļy,	His	Asp 30	Cys	Phe
,	Pro	Val	Leu 35	Phe	Thr	Туг	Asp ;	Ala 40	Ala	Ala	Gly	Met	Leu 45	Ser	Phe	Gly
10	Gly	Arg 50	Leu	Asp	Val	Pro	Lys 55	Gln	Ser	Ser	Gln	Arg 60	Gly	Leu	Thr	Ala
	Arg 65	Glu	Arg	Phe	Gln	Asn 70	Leu	Asp	Lys	Lys	Ala 75	Ser	Ser	Glu	Gly	Gly 80
15	Thr	Ala	Ala	Gly	Ala 85	Gly	Leu	Asp	Ser	Leu ' 90	His	Lys	Asn	Ser	Val 95	Ser
20	Gln	Ile	Ser	Val 100	Leu	Ser	Gly	Gly	Lys 105	Ala	Lys	Cys	Ser	Gln 110	Phe	Cys
	Thr	Thr	Gly 115	Met	Asp	Gly	Gly	Met 120	Ser	Ile	Trp	Asp	Val 125	Lys	Ser	Leu
25	Glu	Ser 130	Ala	Leu	Lys	Asp	Leu 135	Lys	Ile	Lys			,			
30	(2)	INF						NO: 3				•		*		
			(i)	~ (A) L B) T	ENGI YPE :	H: 2 ami	ERIS 3 am no a lin	ino cid		ls	1	ı '			
35			(xi)					PTIO		EQ I	D NO	: 22	8:			
	Leu 1	Gly	Ser	Leu	Ser 5	Thr	Ala	Pro	Ser	Ser 10		Leu	Pro	Thr	Leu 15	Gly
40	Ala	Arg	Arg	Thr 20	·Arg	Ser	Lys									
45	(2)	INF	ORMA	TION	FOR	SEQ	ID :	NO:	229:							
50				((A) I (B) I (D) I	ENGI YPE : OPOI	TH: 1 ami OGY:	ERIS 133 a 100 a 11r	mino cid ear	aci		ı. 22	g.			
	Met	Thr												Ala	Ala	Trp
55	1 Val		Leu	Ala 20			Leu	Gly	Val 25			Tyr	Ala	Ala 30	15 Ala	Val
60	Leu	Leu	Gly 35	Ala		Cys	Ala	Thr 40	Ile		Val	Thr	Ser 45	Leu	Ala	Met

	Thr	Ala 50	Asp	Leu	Ile	Gly	Pro 55	His	Thr	Asn	Ser	Gly 60	Ala	Phe	Val	Туг
5	Gly 65	Ser	Met	Ser	Phe	Leu 70	Asp	Lys	Val	Ala	Asn 75	Gly	Leu	Ala	Val	Met 80
10	Ala	Ile	Gln	Ser	Leu 85	His	Pro	Cys	Pro	Ser 90	Glu	Leu	· Cys	Суз	Arg 95	Ala
	Cys	Val	Ser	Phe 100	Tyr	His	Trp	Ala	Met 105	Val	Ala	Val		Gly 110	Gly	Val
15	Gly	Val	Ala 115	Ala	Ala	Leu	Cys	Leu 120	Cys	Ser	Leu	Leu	Leu 125	Trp	Pro	Thr
••	Arg	Leu 130	Arg	Arg	Xaa											
20		•														
	(2)	INFO	ORMA	NOI	FOR	SEQ	ID N	10: 2	230:							
25				(1	A) Li B) T D) T	ENGT YPE: OPOL	H: 28 amin OGY:	8 am no ja lin	ino a cid ear	acid		. 22/	٦.			
30	01															
50	1	гÀг	PIO	inr	5 5	гуs	Ser	Leu	Pro	Leu 10	Met	Trp	Met	Ile	Leu 15	Met
35	Gln	Pro	Ile	Ile 20	Met	Ile	Ser	Met	Met 25	Ser	Asn	Gly			•	
	(2)	INFO	RMAT	MOI	FOR	SEQ	ID N	ю: 2	31:							
40		((i) S		A) LI	engii		lam	ino a		5					
		((xi)	(I SEQU			OGY:			en er	NO.	231				
1 5	Met 1													Leu	Lys 15	Tyr
50	Leu	Leu	Met	Leu 20	Leu	Cys	Met	Phe	Val 25	Asn	Arg	Gly	Met	Ser 30	Lys	Asp
	Ser	Thr	Lys 35	Lys	Pro	Gly	Gln (Glu 40	Lys	Leu	Lys	Val	Ser 45	Leu	Gly	Ser
55	Ile	Leu 50	Asn	Met :	Lys	Ser	Gln . 55	Arg	Pro	Leu	Ser	Trp 60	Cys			
50	(2)	INFO	RMAT	ION :	FOR	SEQ	ID N	0: 2	32:							

```
(i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 29 amino acids
                     (B) TYPE: amino acid
 5
                     (D) TOPOLOGY: linear
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 232:
      Met Met Glu Arg Ser Met Met Ile Leu Leu Met Ala Ala Ser Met Thr
                                           10
10
      Met Thr Ser Thr Gln Leu Trp Ser Phe Cys Cys Val His
                   20
15
      (2) INFORMATION FOR SEQ ID NO: 233:
             (i) SEQUENCE CHARACTERISTICS:
                     (A) LENGTH: 18 amino acids
20
                     (B) TYPE: amino acid
                     (D) TOPOLOGY: linear
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 233:
      Met Trp Tyr Gln Leu Ala Lys Glu Glu Pro Gly Val Gly Ala Cys Ala
25
                        5 '
                                           10
      Leu Asp
30
      (2) INFORMATION FOR SEQ ID NO: 234:
                                      1
             (i) SEQUENCE CHARACTERISTICS:
35
                     (A) LENGTH: 2 amino acids
                     (B) TYPE: amino acid
                     (D) TOPOLOGY: linear
              (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 234:
40
      Leu Xaa
        1
45
      (2) INFORMATION FOR SEQ ID NO: 235:
              (i) SEQUENCE CHARACTERISTICS:
                     (A) LENGTH: 72 amino acids
                     (B) TYPE: amino acid
50
                    (D) TOPOLOGY: linear
              (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 235:
      Met Leu Ile Cys Arg Leu Val Leu Leu Ala Asp Pro Gly Pro Val Asn
55
      Phe Met Val Arg Leu Phe Val Val Ile Val Met Phe Ala Trp Ser Ile
                   20
      Val Ala Ser Thr Ala Phe Leu Ala Asp Ser Gln Pro Pro Asn Arg Arg
60
                                    40
```

	Ala	Leu 50	Ala	Val	Tyr	Pro	Val 55		Leu	Phe	Туг	Phe 60	Val	lle	Ser	Trp
5	Met 65	Ile	Leu	Thr	Phe	Thr 70	Pro	Gln					. '			
10	(2)	INF	ORMA		i							1		ı		
15				(((A) I B) T D) T	ENGI YPE: OPOL	H: 9 ami OGY:	6 am no a lin		acid		: , 23:	6:			
20	Met 1	Arg	Ser	Leu	Leu 5	Leu	Leu	Ser	Ala	Phe 10	Cys	Leu	Leu	Glu	Ala 15	Ala
	Leu	Ala	Ala	Glu 20	Val	Lys	Lys	Pro	Ala 25	Ala	Ala	Ala	Ala	Pro 30	Gly	Thr
25	Ala	Glu	Lys 35	Leu	Ser	Pro	Lys	Ala 40	Ala	Thr	Leu	Ala	Glu 45	Arg	Xaa	Pro
	Ala	Trp 50	Pro	Ser	Ala	Cys	Thr 55	Arg	Pro	Trp	Pro	Arg 60	Thr	Arg	Gln	Trp
30	Arg 65	Thr	Ser	Trp	Суѕ	His 70	Pro	Trp	Trp	Trp	Pro 75	Arg	Arg	Trp	Gly	Ser 80
35	Cys	Arg	Trp	Ala	Ala 85	Arg	Arg	Pro.	Arg	Arg 90	Arg	Arg	Pro	Arg	Gln 95	Cys
40	(2)	INFO	ORMAT	rion	FOR	SEQ	ID N	10: 2	37:							
45				() ()	A) Li B) T D) T	ENGT YPE: OPOL	H: 14 amii XGY:	43 ar no ao line		acie	_	: 237	':			
50	Met 1	Arg	Ser	Leu	Leu 5	Leu	Leu	Ser	Ala	Phe 10	Cys	Leu	Leu	Glu	Ala 15	Ala
	Leu	Ala ,	Ala	Glu 20	Val	Lys	Lys	Pro	Ala 25	Ala	Ala	Ala	Ala	Pro 30	Gly	Thr
55	Ala	Gl u	Lys 35	Leu	Ser	Pro	Lys	Ala 40	Ala	Thr	Leu	Ala	Glu 45	Arg	Lys	Arg
60	Pro	Gly 50	Leu	Gln	Leu	Val	Pro 55	Gly	His	Gly	Gln	Gly 60	Pro	Gly	Ser	Gly

	Glu 65	His	Pro	Gly	Val	Thr 70	Arg	Gly	Gly	Gly	Leu 75	Val	Ala	'Gly	Ala	Arg 80
5	Val	Ala	Gly	Arg	Gln 85	Gly	Asp	His	Gly	Val 90	Ala	Gly	Gln	Gly	Ser 95	Ala
	Glu	Arg	Arg	Ala 100	Ala	Ala	Arg	Arg	Gly 105	Gly	Ala	Arg	Arg	Pro 110	Gly	Arg
10	Ala	Ala	Ala 115	Leu	Thr	Gln	Gln	Leu 120	His	Gly	Ala	Gln	Arg 125	As p	Lėu	Glu
15	Ala	Gly 130	Gln	Pro	Thr	Val	Arg 135	Thr	Gln	Leu	Ser	Glu 140	Leu	Arg	Xaa	
20	(2)	INF		(ENCE A) L B) T	CHA ENGT YPE:	RACT	ERIS 42 a no a	TICS mino		ds.					
25	Met 1	Arg		SEQ	UENC	E DE	SCRI	PTIO	N: S				8: Leu	Glu	Ala 15	Ala
30		Ala	Ala	Glu 20		Lys	Lys	Pro	Ala 25	Ala	Ala	Ala	Ala	Pro 30	Gly	Thr
	Ala	Glu	Lys 35		Ser	Pro	Lys	Ala 40		Thr	Leu	Ala	. Ġlų 45	Arg	Xaa	Arg
35	Pro	Gly 50		Gln	Leu	Val	Pro 55		His	Gly	Gln	Gly 60	Pro	Gly	Ser	Gly
40	65					70					75		Ala			80
	Val	Ala	Gly	Arg	Gln 85		Asp	His	Gly	Val 90		Gly	Gln	Gly	Ser 95	
45	Glu	Arg	Arg	Ala 100		Ala	Arg	Arg	Gly 105		Ala	Arg	Arg	Pro 110		Arg
	Ala	Ala	Ala 115		Thr	Gln	Gln	Leu 120		Gly	Ala	Glr	125		Leu	Glu
50	Ala	130		Pro	Thr	· Val	135		Gln	Leu	Ser	140	Leu	Arg	Ī	
55	(2)	INF		TION SEQU	JENCI	E CHI	ARACT	TERIS		3:	ds					
60					(B)	TYPE	: am	ino a								

			(xi)	SEX	QUEN	CE DI	ESCRI	PTI	ON: S	SEQ I	D NC): 2 3	89:			
5	As _i	Pro l	o Glu	Ala	Ala		Ser	Gly	Glu	Pro 10		Asn	Lys	Arg	Thr 15	Pro
-	Ası	Let	ı Pro	Glu 20		a Glu	Туг	. Val	Lys 25		Glu ,	Ile	Gln	Glu 30	Asn	Glu
10	Glu	ı Ala	a Val 35		Lys	: Met	Leu	Val		Ala	Thr	Arg	Glu 45		Glu	Glu
	Va]	. Val	Val	Asp	Glu	Ser										•
15													ŧ			
	(2)	INF	ORMA	TION	FOR	SEQ	ID:	NO:	240:							
20				• ((A) I (B) I (D) I	CHA LENGI TYPE: TOPOI LE DE	H: 6 ami OGY:	3 am no a lin	uino cid ear	acid		: 24	0:			
25	Gln 1	Lys	Leu	Lys	Arg 5	Lys	Ala	Glu	Glu i	Asp 10	Pro	Glu	Ala	Ala	Asp 15	Ser
30	Gly	Glu	Pro	Gln 20	Asn	Lys	Arg	Thr	Pro 25	Asp	Leu	Pro	Glu	Glu 30	Glu	Tyr
	Val	Lys	G1u 35	Glu	Ile	Gln	Glu	Asn 40	Glu	Glu	Ala	Val	Lys 45	Lys	Met	Leu
35	Val	Glu 50	Ala	Thr	Arg	Glu	Phe 55	Glu	Ġlu	Val	Val	Val 60	Asp	Glu	Ser	
40	(2)	INF	ORMA:	SEQU:	ENCE A) L		RACT H: 1	ERIS 13 a	rics mino		ds					
45			(xi)	(D) T	OPOL	OGY:	lin	ear	≅Q II	O NO:	24:	l:			
	Lys 1	Ala	Met	Glu	Lys 5	Ser	Ser	Leu	Thr	Gln 10	His	Ser	Trp	Gln	Ser 15	Leu
50	Lys	Asp	Arg	Tyr 20	Leu	Lys	His	Leu	Arg 25	Gly	Gln	Glu	His	Lys 30	Tyr	Leu
55	Leu	Gly	Asp 35	Ala	Pro	Val	Ser	Pro 40	Ser	Ser	Gln	Lys	Leu 45	Lys	Arg	Lys
	Ala	Glu 50	Glu	Asp	Pro	Glu	Ala 55	Ala	Asp	Ser	Gly	Glu 60	Pro	Gln	Asn	Lys

Arg Thr Pro Asp Leu Pro Glu Glu Glu Tyr Val Lys Glu Glu Ile Gln 65 70 75 80

	Glu	Asn	Glu	Glu	Ala .85	Val	Lys	Lys	Met	Leu 90	Val	Glu	Ala	Thr	Arg 95	Glu
5	Phe	Glu	Glu	Val 100	Val	Val	Asp	Glu	Ser 105	Pro	Pro	Asp	Phe	Glu 110	Ile	His
	Ile												•			
10						•										
	(2)	INFO	ORMA'I	rion	FOR	SEQ	ID i	NO: 2	242 :				i	t		
15				~ (. (.	A) L B) T D) T	ENGT YPE: OPOL	H: 1 ami OGY:	ERIS 48 a no a lin PTIO	mino cid ear	aci		: 24				
20	Leu													Ala	Thr	Ile
	1				5					10				_	15	
25	Pro	Leu	Val	Pro 20	Gly	Arg	Asp	Glu	Asp 25	Phe	Val	Gly	Arg	Asp 30	Asp	Phe
	Asp	Asp	Ala 35	Asp	Gln	Leu	Arg	Ile 40	Gly	Asn	Asp	Gly	Ile 45		Met	Leu
30	Thr	Phe 50		Met	Ala	Phe	Leu 55		Asn	Trp	Ile	Gly 60	Phe	Phe	Leu	Ser
35	65	_				70					75					80 Gly
	Phe	Gly	Leu	Ser	Leu 85		Lys	Trp	Ile	Leu 90		· Val	Arg	Phe	Ser 95	Thr
40	Tyr	Phe	Pro	Gly 100		Phe	Asp	Gly	Gln 105		Trp	Leu	Tr	110		Phe
	Leu	Val	Leu 115		Phe	. Leu	Leu	120		Arg	Gly	Phe	11e		Tyr	Ala
45	Lys	Val 130		Lys	Met	Pro	Glu 135		Phe	Ser	Asr	140		Arg	Thr	Arg
50	Val 145		Phe	: Ile	!											
	(2)	INE	ORMA	TION	FOF	SEÇ) ID	NO:	243 :							
55			(i)	_	(A) (B)	LENG TYPE	TH: : am	TERIS 24 au ino	mino acid	aci	ds					
60			(xi					: li: IPTI		SEQ	ID N	0: 2	43 :			

		1	y nu	Y	1 61	5 5	a 11	e se	er G1		e G1 0	y Le	u Se	r, Le		.e Ly .5
5	Tr	p Il	e Le	_	e Va O	l Ar	g Ph	e Se	r				. '			•
10	(2) IN		ATIO SEQ	UENC	Е СН	ARAC	TERI		s:	ds	I		ı	1	
15			(xi) SE	(B) (D)	TYPE TOPO	: am	ino : li	acid near		•	0: 2	44:	•		
	Me	t Ly: 1	s His	s Lei	u Sei	r Ala	а Ттү) Ası	n Phe	Thi 10		s Lei	u Thi	r Phe	e Len 19	u Glr 5
20	Le	u Trį	o Glu	u Ile 20	e Phe	∋ Glı	ı Gly	, Sei	val 25		ı Ası	ı Cys	s Glr	1 Thi 30		ı Thr
25	Sea	г Туз	Ser 35	t Lys	s Lev	Glr	ılle	Lys	: Туг)	Thr	Ph∈	e Sei	Arg	_	/ Sei	Thr
	Ph€	∓ Ту т 50		•	1								ı		r	
30	(2)	INF	ORMA	ATION	I FOR	SEQ	ID	NO:	245:							
35					(A) I (B) 7 (D) 7	LENGI LYPE : LOPOI	TH: 2 ami .OGY:	213 a ino a : lir	amino acid near	aci		n. 24				
40	Phe 1	: Ser				Arg								Arg	Val	Glu
	Ser	Lys	Ala	Thr 20	Ser	Ala	Arg	Cys	Gly 25	Leu	Trp	Gly	Ser	Gly 30	Pro	Arg
45	Arg	Arg	Pro 35	Ala	Ser	Gly	Met	Phe 40	Arg	Gly	Leu	Ser	Ser 45	Trp	Leu	Gly
50	Leu	Gln 50	Gln	Pro	Val	Ala	Gly 55	Gly	Gly	Gln	Pro	Asn 60	Gly	Asp	Ala	Pro
	Pro 65	Glu	Gln	Pro	Ser	Glu 70	Thr	Val	Ala	Glu	Ser 75	Ala	Glu	Glu	Glu	Leu 80
55	Gln	Gln	Ala	Gly	Asp 85	Gln	Glu	Leu	Leu	His 90	Gln	Ala	Lys	Asp	Phe 95	Gly
	Asn	Tyr	Leu	Phe 100	Asn	Phe	Ala	Ser	Ala 105	Ala	Thr	Lys	Lys	Ile 110	Thr	Glu
50	Ser	Val	Ala	Glu	Thr	Ala	Gln	Thr	Ile	Lys	Lys	Ser	Val	Glu	Glu	Gly

			115		•			120		4			125			
5	Lys	Ile 130	Asp	Gly	I le	Ile	Asp 135	Lys	Thr	Ile	Ile	Gly 140	Asp	Phe	Gln	Lys
J	Glu 145	Gln	Lys	Lys	Phe	Val 150	Glu	Glu	Gln	His	Thr 155		Lys	Ser	Glu	Ala 160
10	Ala	Val	Pro	Pro	Trp 165		Asp	Thr	Asn	Asp 170	Gľu	Glu	Thr	Ile	Gln 175	Gln
	Gln	Ile	Leu	Ala 180	Leu	Ser	Ala	Asp	Lys 185	Arg	Asn	Phe	Leu	Arg 190	Asp	Pro
15	Pro	Ala	Gly 195	Val	Gln	Phe	Asn	Phe 200	Asp	Phe	Asp	Gln	Met 205	Tyr	Pro	Val
20	Ala	Leu 210	Val	Met	Leu											
	(2)	INF	ORMA'	rion	FOR	SEQ	ID I	NO:	246:							
25				- (A) I B) I D) I	ENGT YPE: OPOL	H: 4 ami OGY:	9 am no a lin	ear	acid		. 24	<i>c</i> .			
30				_					N:S					**-1	Dh.a	·m
	1				5				ı	10			•		15	Trp
35	Arg	Asn	Tyr	Phe 20	-	Arg	Val	Ser	Leu 25	Ile	Lys	Gln	Ser	Ala 30	Gln	Leu
	Thr	Ala	Leu 35		Ala	Gln	Gln	Gln 40		Ala	Gly	Lys	Gly 45		Glu	Glu
40	Gln				٠											
45	(2)	INF		TION												
			(i)						TICS nino		is					
50			(xi)		(D)	TYPE: TOPOI TE DE	LOGY	: li		SEQ I	ID NO): 24	17:			
55	Ser 1		Ser	Pro	Gly 5		. Ser	Glu	Phe	val 10		Asp	Ala	Phe	Asp 15	Ala
	Cys	Asn	Lev	Asn 20		Glu	a Asp	Leu	Arg 25		Glu	ı Met	: Glu	Glr 30		Val
60	Lev	a Asp	Lys 35	_	Glr	Glu	Glu	1 Thi 40		\Val	Leu	ı Glu	ı Glu 45		Ser	Ala

	Asp	Trp 50	Glu	Lys	Glu	ı Lev	Glr 55		Glu	ı Lev	ı Glr	Glu 60		Glu	ı Val	. Va
5	Thr 65	Glu	Ser	Glu	Lys	Arg 70		Glu	Asn	Trp	75		, ' :			
10	(2)	INF	ORMA	TION	FOR	SEQ	ID	NO:	248:			f				
15				(A) I B) 1 D) 1	E CHA LENGI TYPE: TOPOI TE DE	TH: 6 ami OGY:	52 ar ino a : lir	mino acid near	ació): ₂₄	.8:	,		-
20	1				5					10					Ala 15	
	Cys	Gly	Leu	Trp 20	Gly	Ser	Gly	' Pro	Arg 25		Arg	Pro	Ala	Ser 30	Gly	Met
25	Phe	Arg	Gly 35	Leu	Ser	Ser	Trp	Leu 40	Gly	Leu	Gln	Gln	Pro 45	Val	Ala	Gly
20	Gly	Gly 50	Gln	Pro	Ąsn	Gly	Asp 55	Ala	Pro	Pro	Glu	Gln 60	Pro	Ser		
30	(2)	TNEC	חגאפר	TON	EOD.	œ	TD 1	170	0.40					·		
35	(2)		(i) s	EEQUI () ()	ENCE A) L B) T D) T	CHAI ENGT YPE: OPOL E DE:	RACT H: 6 ami OGY:	ERIS 5 am no a lin	TICS ino cid ear	acid		: 24	9:			
40	Pro 1	Val	Ala	Gly	Gly 5	Gly	Gln	Pro	Asn	Gly 10	Asp	Ala	Pro	Pro	Glu 15	Gln
45	Pro :	Ser	Glu	Thr 20	Val	Ala	Glu	Ser	Ala 25	Glu	Glu	Glu	Leu	Gln 30	Gln	Ala
	Gly	Asp	Gln 35	Glu	Leu	Leu	His	Gln 40	Ala	Lys	Asp	Phe	Gly 45	Asn	Tyr	Leu
50	Phe i	Asn 50	Phe	Ala	Ser	Ala	Ala 55	Thr	Lys	Lys	Ile	Thr 60	Glu	Ser	Val	Ala
55	Glu 65															
	(2)	INFO	RMAT	ION	FOR	SEQ	ID N	ю: 2	50:							
50		(i) S			CHAF ENGTI					s					

				(D) T	OPOL	amı OGY:	lin	ear				•			
			(X1)	SEQ	UENC	E DE	SCRI	Prio	N: S	EQ I	D NO	: 25	Ų : <i>i</i>		٠	
5	Phe 1	Gln	Lys	Glu	Gln 5	Lys	Lys	Phe	Val	Glu 10	Glu	Gln	·His	Thr	Lys 15	Lys
10	Ser	Glu	Ala	Ala 20	Val	Pro	Pro	Trp	Val 25	Asp	Thr	Asn I	Asp	Glu 30	Glu ,	Thr
10	Ile	Gln	Gln 35	Gln	Ile	Leu	Ala	Leu 40	Ser	Ala	Asp	Lys	Arg 45	^l Asn	Phe	Leu
15	Arg	Asp 50	Pro	Pro	Ala	Gly	Va'l 55		Phe	Asn	Phe	Asp 60	Phe	Asp	Gln	Met
	Tyr 65	Pro	Val	Ala	Leu	Val 70	Met	Leu				•				٠
20							,				1					
	(2)	INF	ORMA!	PION	FOR	SEQ	ID I	NO:	251:							
25				(A) L B) T D) T	ENGT YPE: OPOL	ami OGY:	8 am no a lin	ino cid ear	: acid EQ I		: 25	1:	,		
30	Pro 1	Phe	Ile	Cys	Val 5	Ala	Arg	Asn	Pro	Val 10	Ser	Arg	Asn	Phe	Ser 15	Ser
35	Pro	Ile	Leu	Ala 20	Arg	Lys	Leu	Cys	`Glu 25	Gly	Ala	Ala	1			
	(2)	INF	ORMA'	TION	FOR	SEQ	ID :	NO:	252:							
40			(i)	(A) I	ENGI	RACT H: 3	3 an	nino	: acid	s					
45			(xi)				OGY: SCRI			EQ I	D NO	: 25	2:			
73	Lys 1	Glu	Asp	Pro	Ala 5	Asn	Thr	Val	Tyr	Ser 10	Thr	Val	Glu	Ile	Pro 15	Lys
50	Lys	Met	Glu	Asn 20	Pro	His	Ser	Leu	Leu 25	Thr	Met	Pro	Asp	Thr 30	Pro	Arg
	Leu															
55																
	(2)	INF	ORMA	TION	FOR	SEQ	ID	NO:	253:							
60			(i)				RACT		_	: o aci	ds					

			(xi)		D) I	YPE: OPOL E DE	OGY :	lin	ear	EQ I	d No	: 25	3:			
5	Ala 1	Ser	Ala	Val	Leu 5	Leu	Asp	Leu	Pro	Asn 10	Ser	Gly	Gly	Glu	Ala 15	Gl r
10	Ala	Lys	Lys	Leu 20	Gly	Asn	Asn	Cys	Val 25	Phe	Ala	Pro	Ala	Asp 30	Val	Thr
	Ser	Glu	Lys 35	Asp	Val	Gln	Thr	Ala 40	Leu	Ala	Leu	Ala	Lys 45		Lys	. Phe
15	Gly	Arg 50	Val	Asp	Val	Ala	Val 55	Asn	Cys	Ala	Gly	Ile 60	Ala	Val	Ala	Ser
	Lys 65	Thr	Tyr	Asn	Leu	Lys 70	Lys	Gly	Gln	Thr	His 75	Thr	Leu	Glu	Asp	Phe 80
20	Gln	Arg	Val	Leu	Asp 85	Val	Asn	Leu	Met	Gly 90	Thr	Phe	Asn	Val	Ile 95	Arg
25	Leu	Val	Ala	Gly 100	Glu	Met	Gly	Gln	Asn 105	Glu	Pro	Asp	Gln	Gly 110	Gly	Gln
	Arg	Gly	Val 115	Ile	Ile	Asn	Thr	Ala 120	Ser	Val	Ala	Ala	Phe 125	Glu	Gly	Gln
30	Val	Gly 130	Gln	Ala	Ala	Tyr	Ser 135	Ala	Ser	Lys	Gly	Gly 140	Ile	Val	Gly	Met
	Thr 145	Leu	Pro	Ile	Ala	Arg 150	Asp	Leu	Ala	Pro	Ile 155	Gly	Ile	Arg	Val	Met 160
35	Thr	Ile	Ala	Pro	Gly 165	Leu	Phe	Gly	Thr	Pro 170	Leu	Leu	Thr	Ser	Leu 175	Pro
40	Glu	Lys	Val	Cys 180	Asn	Phe	Leu	Ala	Ser 185	Gln	Val	Pro	Phe	Pro 190	Ser	Arg
	Leu	Gly	Asp 195	Pro	Ala	Glu	Tyr	Ala 200	His	Leu	Val	Gln	Ala 205	Ile	Ile	Glu
45	Asn	Pro 210	Phe	Leu	Asn	Gly	Glu 215	Val	Ile	Arg	Leu	Asp 220	Gly	Ala	Ile	Arg
	Met 225	Gln	Pro													
50				٠												
	(2)			MOI		_							•			
55				(1 (1	A) L: B) T D) T	ENGT YPE: OPOL	H: 2: ami OGY:	9 am no ao line	ino a cid ear	acid						
60	Ser			SEQU Ala										ጥህን፦	Ser	Δ 1 =

	1	5	10	15
5	Ser Lys G	ly Gly Ile Val Gl	y Met Thr Leu Pro Il 25	e Ala
	(2) INFOR	MATION FOR SEQ ID	NO: 255:	
10		(B) TYPE: an (D) TOPOLOGY	61 amino acids mino acid Y: linear	
15		-	RIPTION: SEQ ID NO: 2 u Leu Ala Trp Asp Ty 10	
20	Ala Gln L	ys His Lys Asn Tr 20	p Arg Phe Gln Lys Th 25	ur Arg Gln Thr Trp 30
		eu His Met Tyr As 35	p Ser Asp Lys Val Pr 40	o Asp Glu His Phe 45
25	Ser Thr L	-	eu Glu Çly Leu Gln Gl 5 6	y Arg O
30		MATION FOR SEQ ID .) SEQUENCE CHARAC (A) LENGTH: (B) TYPE: an (D) TOPOLOGY	TERISTICS: 22 amino acids nino acid	
35		i) SEQUENCE DESCR	RIPTION: SEQ ID NO: 2	
40	1 Ile Asn L	5 ys Leu Cys Phe 20	10	15
45	(2) INFOR	MATION FOR SEQ ID) NO: 257:	
	(i		22 amino acids	
50	к)	(B) TYPE: ar (D) TOPOLOG (i) SEQUENCE DESCI		257 :
55	1	le Lys Tyr Cys Le 5 .sp Asn Ile Gly	eu Thr Leu Met Gln As 10	sn Ala Gln Leu Ser 15
60		20		

	(2) INFORMATION FOR SEQ ID NO: 258:
5	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 25 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 258:
10	Lys Val Ser Tyr Leu Arg Pro Leu Asp Phe Glu Glu Ala Arg Glu Leu 1 5 10 15
	Phe Leu Leu Gly Gln His Tyr Val Phe 20 25
15	
	(2) INFORMATION FOR SEQ ID NO: 259:
20	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 25 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 259:
25	Met Glu Arg Arg Cys Lys Met His Lys Arg Xaa Ile Ala Met Leu Glu 1 5 10 15
30	Pro Leu Thr Val Asp Leu Asn Pro Gln 20 25
	(2) INFORMATION FOR SEQ ID NO: 260:
35	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 23 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear
40	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 260:
	Ser His Ile Val Lys Lys Ile Asn Asn Leu Asn Lys Ser Ala Leu Lys 1 10 15
45	Tyr Tyr Gln Leu Phe Leu Asp 20
50	(2) INFORMATION FOR SEQ ID NO: 261:
	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 64 amino acids (B) TYPE: amino acid`
55	(D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 261:
	Phe Thr His Leu Ser Thr Cys Leu Leu Ser Leu Leu Leu Val Arg Met 1 5 10 15
60	Ser Gly Phe Leu Leu Ala Arg Ala Ser Pro Ser Tle Cvs Ala Leu

				20					25					, 30		
5	Asp	Ser	Ser 35	Cys	Phe	Val	Gln	Glu 40	Tyr	Cys	Ser	Ser	Tyr 45	Ser	Ser	Ser
5	Cys	Phe 50	Leu	His	Gln	His	Phe 55	Pro	Ser	Leu	Leu	Asp 60	His	Leu	Cys	Gln
10						t		·				ı		ı	•	
15	(2)			PION:				ı			•		•			
			(i) :	(A) L B) T	ENGT YPE:	RACT H: 2 ami OGY:	3 am no a	ino cid		s					-
20			(xi)	SEQ	UENC	E DE	SCRI	PTIO	N: S	EQ I	D NO	: 26	2:			
	Phe 1	Leu	Leu	Leu	Ala 5	Arg	Ala	Ser	Pro	Ser 10	Ile	Cys	Ala	Leu	Asp 15	Ser
25	Ser	Cys	Phe	Val 20	Gln 	Glu	Tyr			•			•	,	,	
30	(2)	INFO	ORMA!	rion	FOR	SEQ	ID I	NO: :	263:					•		
			(i)	-	A) I	ENGI	RACT H: 5	3 am	ino		s		1 1			
35			(xi)		D) 1	OPOL	OGY:	lin	ear	EQ I	D NO	: 26	3:			
40	Pro 1	Asp	Gly	Arg	Val 5		Asn	Ile	Pro	Gln 10	Gly	Met	Val	Thr	Asp 15	Gln
	Phe	Gly	Met	Ile 20	Gly	Leu	Leu	Thr	Phe 25		Arg	Ala	Ala	Glu 30	Thr	Asp
45	Pro	Gly	Met 35	Val	His	Leu	Ala	Leu 40	Gly	Ser	Asp	Leu	Thr 45	Thr	Leu	Gly
	Leu	Asn 50	Leu	Asn	Ser											
50				٠					•							
	(2)	INF	ORMA'	TION	FOR	SEQ	ID	NO:	264:							
55				((A) I (B) T (D) T	ENGI TYPE : TOPOI	TH: 4 ami OGY:	l am no a lir	nino Icid Near	acid): 26	. 4 -			
														7 cm	Wal.	Leu

	1		5		10		15
5	Gln L		la Ala Val 20 .	l Glu Leu	Phe Asn Ar	g Asp Trp	Arg Tyr His
	Lys G	ilu Glu Ai 35	g Val Tr	o Ile Thr 40	Arg		
10	(2) I	NFORMATIO	ON FOR SEC	ID NO:	265:		
15			(B) TYPE:	TH: 24 am : amino a LOGY: lin	ino acids cid): D: 265:	
20	Val H 1	is Leu Al	a Leu Gly 5	Ser Asp	Leu Thr Thr	Leu Gly	Leu Asn Leu 15
	Asn S		u Asn Leu O	Tyr Pro			
25					,		
	(2) II	NFORMATIO	n for seq	ID NO: 2	¹ 2 6 6 :		
30			(B) TYPE: (D) TOPOI	TH: 41 am amino a OGY: lin	ino acids cid	D: 266:	
35	His As 1				Pro Gly Ser	•	
40	(2) II	NFORMATIO	n for seq	ID NO: 2	267 :		
45			(B) TYPE:	H: 75 am amino ao OGY: line	ino acids cid	o: 267:	
50	Gly Ai 1	rg Ile Il	e Asp Thr	Ser Leu	Thr Arg Asp 10	Pro Leu V	Val Ile Glu 15
	Leu Gl	ly Gln Ly: 2		Ile Pro	Gly Leu Glu 25	Gln Ser 1	Leu Leu Asp 30
55	Met Cy	ys Val Gly 35	y Glu Lys	Arg Arg 40	Ala Ile Ile	Pro Ser I	His Leu Ala
	Tyr G	ly Lys Ar	g Gly Phe	Pro Pro 55	Ser Val Pro	Ala Asp 1 60	Ala Val Val
60	Gln Ty	r Asp Va	l Glu Leu	Ile Ala	Leu Ile Arg		

					70					25					
	65				70					75					
5	(2) IN	FORMAT	'ION I	FOR :	SEQ	ID N	0: 2	68:				,			
10			(Þ	A) LE B) TY D) TO	NGTH PE: POLC	H: 16 amir XGY:	ami no ac line	ino a cid ear	cids		268	3: , ¹		•	
15	Ile Hi 1	s Tyr	Thr	Gly 5	Ser	Leu	Val	Asp	Gly 10	Arg	Ile	Ile	Asp	Thr 15	Ser
20	(2) IN	IFORMA'I	rion :	FOR	SEQ	ID N	ю: 2	69:		,					
25			(E	A) LE 3) T'S 2) TC	ENGTI (PE : OPOL(H: 20 amii OGY:	o ami no ac line	ino a cid ear	acid		: 269	· 9:			*
30	Cys Gl 1	.u Ser	Pro	Glu 5	Ser	Pro	Ala	Gln	Pro 10	Ser	Gly	Ser	Ser	Leu 15	Pro
25	Ala Tr	р Туг	His 20				,					','		٠	
35	(2) IN	VFORMA	rion	FOR	SEQ	ID N	ю: 2	270:							
40			(1	A) Li 3) T 0) T	engt YPE: OPOL	H: 9 ami: OGY:	5 am no a lin	ino a cid ear	acid		: 27	0:			
45	Glu Gl	lu Ala	Gly	Ala 5	Gly	Arg	Arg	Cys	Ser 10	His	Gly	Gly	Ala	Arg 15	Pro
50	Ala G	ly Leu	Gly 20	Asn	Glu	Gly	Leu	Gly 25	Leu	Gly	Gly	Asp	Pro 30	Asp	His
	Thr As	sp Thr 35		Ser	Arg	Ser	Lys 40	Gln	Arg	Ile	Asn	Asn 45	Trp	Lys	Gl
55		ys His 50	Lys	Val	Ile	Met 55	Ala	Ser	Ala	Ser	Ala 60	Arg	Gly	Asn	Glı

Asp Lys Asp Ala His Phe Pro Pro Pro Ser Lys Gln Ser Leu Leu Phe 65 70 75 80

Cys Pro Lys Ser Lys Leu His Ile His Arg Ala Glu Ile Ser Lys

85 90, 95 5 (2) INFORMATION FOR SEQ ID NO: 271: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 23 amino acids' (B) TYPE: amino acid 10 (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 271: Ser Lys Gln Arg Ile Asn Asn Trp Lys Glu Ser Lys His Lys Val Ile 5 10 15 Met Ala Ser Ala Ser Ala Arg 20 20 (2) INFORMATION FOR SEQ ID NO: 272: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 32 amino acids 25 (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 272: Leu Phe His Trp Ala Cys Leu Asn Glu Arg Ala Ala Gln Leu Pro Arg 30 5 Asn Thr Ala Xaa Ala Gly Tyr Gln Cys Pro Ser Cys Asn Gly Pro Ser , 25 35 40 (2) INFORMATION FOR SEQ ID NO: 273: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 185 amino acids (B) TYPE: amino acid 45 (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 273: Phe Tyr Ile Tyr Tyr Arg Pro Thr Asp Ser Asp Asn Asp Ser Asp Tyr 50 Lys Lys Asp Met Val Glu Gly Asp Lys Tyr Trp His Ser Ile Ser His 20 Leu Gln Pro Glu Thr Ser Tyr Asp Ile Lys Met Gln Cys Phe Asn Glu 55

Gly Gly Glu Ser Glu Phe Ser Asn Val Met Ile Cys Glu Thr Lys Ala

Arg Lys Ser Ser Gly Gln Pro Gly Arg Leu Pro Pro Pro Thr Leu Ala

50

	65					70				•	75					80
<u>.</u>	Pro	Pro	Gln	Pro	85 Pro	Leu	Pro	Glu	Thr	Ile 90	Glu	Arg	Pro	Val	Gly 95'	Thr
5	Gly	Ala	Met	Val 100	Ala	Arg	Ser	Ser	Asp 105	Leu	Pro	Tyr	Leu	Ile 110	Val	Gly
10	Val	Val	Leu 115	Gly	Ser	Ile ,	Val	Leu 120	Ile	Ile	Val	Thr	Phe 125	Ile	Pro	Phe
	Суѕ	Leu 130	Trp	Arg	Ala	Trp	Ser 135	Lys	Gln	Lys	His	Thr 140	Thr	Asp	Leu	Gly
15	Phe 145	Pro	Arg	Ser	Ala	Leu 150	Pro	Pro	Ser	Cys	Pro 155	Tyr	Thr	Met	Val	Pro 160
20	Leu	Gly	Gly	Leu	Pro 165	Gly	His	Gln	Ala	Val 170	Asp	Ser	Pro	Thr	Ser 175	Val
20	Ala	Ser	Val	Asp 180	Gly	Pro	Val	Leu	Met 185							
25	(2)	INF	orma	TION	FOR	SEQ	ID :	NO:	2 [/] 74:							
			(i)	SEQU	ENCE	CHA	RACT	ERIS	TICS	:						
30					(A) I (B) T					ació	ls					
30					r (d)	OPOI	OGY:									•
			(xi)	SEC	UENC	E DE	SCRI	PTIC	NN: S	EQ 1	D NC): 27	4:			
35	Tyr 1					Pro			1		Asn		•	Asp	Tyr 15	Lys
35	1		туг	Туг	Arg 5	Pro	Thr	Asp	Ser	Asp 10	Asn	Asp	Ser		15	Lys Leu
35 40	1 Lys	Asr	Tyr Met	Tyr Val 20	Arg 5	Pro Gly	Thr Asp	Asp Lys	Ser Tyr 25	Asp 10	Asn His	Asp Ser	Ser	Ser 30	15 His	
40	1 Lys Gln	Asp	Met Glu 35	Val	Arg 5 Glu	Pro Gly Tyr	Thr Asp	Asp Lys Ile 40	Ser Tyr 25	Asp 10 Trp	Asn His	Asp Ser	Ser Ile Phe 45	Ser 30 Asn	15 His Glu	Leu
	Lys Gln Gly	Asp Pro Glu 50 Ser	Met O Glu 35	Val	Arg 5 Glu	Pro Gly Tyr	Thr Asp Asp	Asp Lys Ile 40	Ser Tyr 25	Asp 10 Trp	Asn His	Asp Ser Cys	Ser Ile Phe 45	Ser 30 Asn	15 His Glu	Leu Gly
40	Lys Glm Gly Lys 65	Asp Pro Glu 50 Ser	Met Glu 35	Tyr 20 1 Thr	Arg 5 Glu	Gly Tyr	Asp Asp Asr 55	Asp Lys Ile 40 Val	Ser Tyr 25 Lys	Asp 10	Asn His	Asp Ser Cys	Ser Ile Phe 45	Ser 30 Asn	15 His Glu	Leu Gly
40 45 50	Lys Glm Gly Lys 65	Asp Pro Glu 50 Ser	Met Glu 35 Ser FORM	Tyr Val 20 Thr Glu	Arg 5 Glu Ser Phe	Pro Gly Tyr Ser	Asp Asp 55	Asp Lys Ile 40 Val	Ser Tyr 25 Lys . Met	Asp 10 Trp Met	Asn His	Asp Ser Cys	Ser Ile Phe 45	Ser 30 Asn	15 His Glu	Leu Gly
40 45	Lys Glm Gly Lys 65	Asp Pro Glu 50 Ser	Met Glu Ser FORMU	Tyr Val 20 Thr Glu	Arg 5 Glu Ser Phe Phe VFOR	Gly Tyr Ser CHLENG	Asp Asp Asr SS ID ARAC TH: am LOGY	NO:	Ser Tyr 25 Lys Met	Asp 10 Trr	Asn His Glr Cys	Asp Cys Glu	Ile Phe 45	Ser 30 Asn	15 His Glu	Leu Gly
40 45 50	Lys Glm Gly Lys 65	Asp Glu Pro 50 50 INI	Met Met Glu 35 Ser (i) (xi	Tyr Val 20 Thr Glu SEQ	Arg 5 Glu Ser Phe Phe (A) (B) (C) (QUEN	Pro Gly Tyr Ser CE SEC	Asp Asp Ass SS ID ARAC TH: am LOGY ESCR	Lys Lys Ile 40 Val Val 30 a ino : li IPTI	Ser Tyr 25 Lys Met 275: STIC: mino acid near ON:	Asp 10 Trr Met	Asn His Glr Cys	Asp Cys Glu 60	Ser Ile 45 45 Thr	Ser 30	His Glu Ala	Leu Gly

	Thi	r Al	a Ly:	s Pho 20	e Ası O	n Asr	n Ası	ı Lys	Arg 25		Asr	ı Le	ı Sei	r Leu 30		
5													,			
	(2)	IN	FORM	TION	1 FOF	Seç	O ID	NO:	276:							
10					UENCE (A) 1 (B) 5 (D) 5	LENG IYPE IOPOI	IH: : am: LOGY	185 a ino a : lir	amino acid near	o aci): 27	76: '	i	·	
15	Asn 1	Thu	r Asn											n Phe	Gln 15	
20	Phe	Ala	a Leu	Asn 20	His	Gln	Lys	Asp	Ile 25	Gln	Val	Leu	Met	: Gly :30		Leu
	Val	Тух	35	Arg	Gln	Gly	Ile	Glu 40	Asn	Ser	Pro	Tyr	Val 45		Leu	Leu
25	Asp	Ala 50	Asn	Gln	Trp	Ala	Asp 55	Ile	Cys	Asp	Ile	Phe 60	Thr	Arg	Asp	Ala
	Суз 65	Ala	Leu	Leu	Gly	Leu 70	Ser	Val	Glu	Ser	Pro 75	Leu	Ser	Val	Ser	Phe 80
30	Ser	Ala	Gly	Cys	Val 85	Ala	Leu	Pro	Ala	Leu 90	Ile	Asn	Ile	Lys	Ala 95	Val
35	Ile	Glu	Gln	Arg 100	Gln	Cys	Thr	Gly	Val 105	Trp	Asn	Gln	Lys	Asp 110	Glu	Leu
	Pro	Ile	Glu 115	Val	Asp	Leu	Gly	Lys 120	Lys	Cys	Trp	Tyr	His 125	Ser	Ile	Phe
40	Ala	Cys 130	Pro	Ile	Leu	Arg	Gln 135	Gln	Thr	Thr	Asp	Asn 140	Asn	Pro	Pro	Met
	Lys 145	Leu	Val	Cys	Gly	His 150	Ile	Ile	Ser	Arg	Asp 155	Ala	Leu	Asn	Lys	Met 160
45	Phe	Asn	Gly	Ser	Lys 165	Leu	Lys	Cys	Pro	Tyr 170	Суз	Pro	Met	Glu	Gln 175	Ser
50	Pro	Gly	Asp	Ala 180	Lys	Gln	Ile	Phe	Phe 185							
	(2)	INFO	ORMAT	NOI	FOR	SEQ	ID N	io: 2	77 :							
55			(i) s	() ()	A) LE 3) TY	NGTI PE:	i: 69 amir	ami no ac	no a		:					
60			(xi)		D) TC					Q ID	NO:	277	':			

	Ser 1	Tyr	Leu	Ser	Ala 5	Cys	Phe	Ala	Gly	Суs 10	Asn	Ser	Thr'	Asn	Leu 15	Thr
5	Gly	Cys	Ala	Cys 20	Leu	Thr	Thr	Val	Pro 25	Ala	Glu	Asn [']	Åla	Thr 30	Val	Val
	Pro	Gly	Lys 35	Cys	Pro	Ser	Pro	Gly 40	Cys	Ģln	Glu	Ala '	Phe 45	Leu	Thr	Phe
10	Leu	Cys 50	Val	Met	Cys	ıle	Сув 55	Ser	Leu	Ile	Gly	Ala 60	Met	Ala	Arg	His
15	Pro 65						•	I		ì						
20	(2)	INF		SEQU	ENCE	CHA	RACI H: 8	NO: S ERIS 34 am	TICS iino		ls			,		
25				SEÇ	UENC : Ile	E DE	SCRI	lir PTIC Arg	N:S		Ser		t	Leu	Lys 15	Ser
30	Туг		. Leu	Gly 20			Phe	e Leu	Leu 25	Leu		Leu	Leu	Gly 30		e Ile
	Pro	Pro	Pro 35		ı Ile	Phe	Gly	7 Ala 40		Ile	Asp	Ser	Thr 45		Lev	Phe
35	Tr	Sei 50		Phe	e Cys	Gly	7 Glu 59		Gly	Ala	Cys	Val		Tyr	Asp) Asn
40	69	5		Arg		: L et 7(r Val	Ser	: Ile	Ala 75		e Ala	. Leu	Lys	s Ser 80
45	(2) IN	FORM	ATIO	N FO	r sea	O ID	NO:	279:	:						
50					(A) (B) (D)	LENG TYPE TOPO	TH: : an LOGY	TERI 182 nino 7: li	amin acid near	o ac		0: 2	79:			
55		n Se 1	r Le	u Ph		r Ar 5	g Ph	e Va	l Ar	g Va		y Va	l Pr	o Thi	va 1	l Asp 5
	Le	u As	p Al		n Gl	y Ar	g Al	a Ar	g Al 2		r Le	u Cy	s Xa	a Xaa		r Asn
60	m~	A+	-a Th	r Tau	re Ne	n Te	u G1	v As	n Le	u Pr	o Hi	s Va	ı Gl	n Le	ı Le	u Pro

	35		40		45	
5	Glu Phe Ser 50	Thr Ala	Asn Ala Gly 55	Leu Leu Tyr	Asp Phe Gln 60	Leu Il
	Asn Val Glu 65	Asp Phe	Gln Gly Val 70	Gly Glu Ser	Glu Pro Asn	Pro Ty
10	Phe Tyr Gln	Asn Leu 85	Gly Glu Ala	Glu Tyr Val 90	Val Ala Leu	Phe Me
	Tyr Met Cys	Leu Leu 100	Gly Tyr Pro	Ala Asp Lys 105	Ile Ser Ile : 110	Leu Th
15	Thr Tyr Asn 115	Gly Gln	Lys His Leu 120	Ile Arg Asp	Ile Ile Asn 1	Arg Arg
20	130		135		Lys Val Thr 1	
	142	•	150	155	Leu Leu Ser 1	160
25		165		Arg Asp Val 170	Arg Arg Leu \	Val Val 175
30	Ala Met Ser	Arg Ala 1 180	Arg			
30	(2) INFORMAT	TON FOR S	SEQ ID NO: 2	80:		٠
35		(A) LEI (B) TYI (D) TOI	CHARACTERIST NGTH: 77 ami PE: amino ac POLOGY: line DESCRIPTION	no acids id	280:	
40	Leu Val Lys 1	Glu Ala L 5	ys Ile Ile i	Ala Met Thr	Cys Thr His A	la Ala 15
45	Leu Lys Arg	His Asp L 20	eu Val Lys 1	Leu Gly Phe 1 25	Lys Tyr Asp A 30	sn Ile
	Leu Met Glu 35	Glu Ala A	la Gln Ile I 40	Leu Glu Ile (Glu Thr Phe I 45	le Pro
50	Leu Leu Leu 6 50	Gln Asn P	ro Gln Asp (55	Gly Phe Ser A	Arg Leu Lys A 60	rg Trp
55	Ile Met Ile 6		is His Gln I 70	eu Pro Pro V 75	/al Ile	
	(2) INFORMAT	ION FOR SI	EQ ID NO: 28	1:		
60	(i) S		HARACTERISTI GTH: 125 ami			

						YPE: OPOL										
			(xi)			E DES				EQ II	ON C	281	L:		,	
5	Asp 1	Thr	Tyr	Pro	Asn 5	Glu	Glu	Lys	Gln	Gln 10	Glu	Arg	Val	Phe	Pro 15	Хаа
10	Xaa	Ser	Ala	Met 20	Val	Asn	Asn	Gly	Ser 25	Leu	Ser	Tyr	Asp	His 30	Glu	Arg
10	Asp	Gly	Arg 35	Pro	Thr	Glu	Leu	Gly 40	Gly	Cys	Xaa	Ala	Ile 45	.Val	Arg	Asn
15	Leu	His 50	Tyr	Asp	Thr	Phe	Leu 55	Val	Ile	Arg	Tyr	Val 60	Lys	Arg	His	Leu
	Thr 65	Ile	Met	Met	Asp	Ile 70	Asp	Gly	Lys	His	Glu 75	Trp	Arg	Asp	Cys	Ile 80
20	Glu	Val	Pro	Gly	Val 85	Arg	Leu	Pro	Arg	Gly 90	Tyr	Tyr	Phe	Gly	Thr 95	Ser
25	Ser	Ile	Thr	Gly 100	Asp	Leu	Ser	Asp	Asn 105		Asp	Val	Ile	Ser 110	Leu	Lys
	Leu	Phe	Glu 115		Thr	Val	Glu	Arg 120	Thr	Pro	Glu	Glu	Glu 125			
30	(2)	INF	ORMA	TION	FOR	SEQ	ID :	NO:	282:							
35				((A) I (B) 7 (D) 7	CHA LENGI TYPE: TOPOL CE DE	H: 8 ami OGY:	35 an ino a : lir	nino acid near	ació		: 28	2:			
40	Leu 1		Arg	Glu	His		Leu	Ser	Lys	Pro 10		Gln	Gly	Val	Gly 15	Thr
	Gly	Ser	Ser	Ser 20		Trp	Asn	Leu	Met 25		Asn	Ala	Met	Val 30		Thr
45	Gln	Тут	: Ile 35		Leu	Thr	Pro	Asp 40		Gln	Ser	Lys	Gln 45		Ala	Leu
50	Trp	Asr 50		r Val	Pro	Cys	Phe 55		Arg	, Asp	Trp	60 60		Gln	Val	His
	Phe 65		: Ile	e His	: G1y	70		, Lys	Lys	a Asr	Leu 75		Gly	Asp	Gly	Leu 80
55	Ala	ı Ile	e Trp	тут	Thr 85											
60	(2)	INI	FORM	ATION	1 FOI	R SEÇ) ID	NO:	283	:						

	(1) SEQUENCE CHARACTERISTICS:
	(A) LENGTH: 32 amino acids
	(B) TYPE: amino acid
_	(D) TOPOLOGY: linear
5	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 283:
	\cdot
	Pro Gly Thr Leu Gln Cys Ser Ala Leu His His Asp Pro Gly Cys Ala
	1 5 10 15
	;
10	Asn Cys Ser Arg Phe Cys Arg Asp Cys Ser Pro Pro Ala Cys Gln Cys
	20 25 30
· 15	
	. 1
	(2) INFORMATION FOR CEO ID NO. 204
	(2) INFORMATION FOR SEQ ID NO: 284:
20	(i) Charmage and a comment
20	(i) SEQUENCE CHARACTERISTICS:
•	(A) LENGTH: 27 amino acids
	(B) TYPE: amino acid
	(D) TOPOLOGY: linear
25	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 284:
23	
	Phe Leu Tyr Asp Val Leu Met Xaa His Glu Ala Val Met Arg Thr His
	1 10 15
20	Gln Ile Gln Leu Pro Asp Pro Glu Phe Pro Ser
30	20 25
	(2) INFORMATION FOR SEQ ID NO: 285:
35	
	(i) SEQUENCE CHARACTERISTICS:
	(A) LENGTH: 6 amino acids
	(B) TYPE: amino acid
	(D) TOPOLOGY: linear
40	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 285:
	Total Total Day ID No. 203.
	Gly Trp Tyr Trp Cys Gly
	1 5
	-
45	
	(2) THEORMATION FOR GEO TO NO. 200
	(2) INFORMATION FOR SEQ ID NO: 286:
	/i/ CECTENION OTTO DE COMPRE COMP
50	(i) SEQUENCE CHARACTERISTICS:
50	(A) LENGTH: 129 amino acids
	(B) TYPE: amino acid
	(D) TOPOLOGY: linear
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 286:
E E	
55	Met Lys Val Gly Ala Arg Ile Arg Val Lys Met Ser Val Asn Lys Ala
	1 5 10 15
	His Pro Val Val Ser Thr His Trp Arg Trp Pro Ala Glu Trp Pro Gln
	20 25 30
60	

	Met	Phe	Leu 35	His	Leu	Ala	Gln	Glu 40	Pro	Arg	Thr	Glu	Vaľ 45	Lys	Ser	Arg
5	Pro	Leu 50	Gly	Leu	Ala	Gly	Phe 55	Ile	Arg	Gln	Asp	Ser 60	Lys	Thr	Arg	Lys
	Pro 65	Leu	Glu	Gln	Glu	Thr 70	Ile	Met	Ser	Ala	Ala 75	Asp	Thr	Ala	Leu	Trp 80
10	Pro	Tyr	Gly	His	Gly 85	Asn	Arg	Glu	His	Gln 90	Glu	Asn	Glu	Leu	Gln 95	Lys
15	Tyr	Leu	Gln	Tyr 100	Lys	Asp	Met	His	Leu 105	Leu	Asp	Şer	Gly	Gln 110	Ser	Leu
	Gly	His	Thr 115	His	Thr	Leu	Gln	Gly 120	Ser	Hís	Asn	Leu	Thr. 125	Ala	Leu	Asn
20	Ile						ě				ı	-				
25	(2)	INFO	ORMA'	rion	FOR	SEQ	ID 1	NO: 2	287:	•		•				4
25			(i) :	(.	A) L	engt	H: 4	9 am	ino .		s		'	÷		
30			(xi)	C	D) T	OPOL	ami: OGY: SCRI	lin	ear	EQ II	ОИО	: 28	7:	•		
•	Ser 1	Leu	His	Lys	Asn 5	Ser	Val	Ser	Gln	Ile 10	Ser	Val	Leu	Ser	Gly 15	Gly
35	Lys	Ala	Lys	Cys 20	Ser	Gln	Phe	Cys	Thr 25	Thr	Gly	Met	Asp	Gly 30	Gly	Met
40	Ser	Ile	Trp 35	Asp	Val	Lys	Ser	Leu 40	Glu	Ser	Ala	Leu	Lys 45	Asp	Leu	Lys
	Ile															
45	(2)	INFO	RMAT	TION	FOR	SEQ	ID N	Ю: 2	288:							
50			(i) s (xi)	· (1 · (1 (1	A) L B) T D) T	ENGT YPE : OPOLA	H: 2: ami OGY:	1 am no a lin	ino a cid ear	acid	_	: 288	3:			
55	Glu 1	Ala	Ser	Lys	Ser 5	Ser	His	Ala	Gly	Leu 10	Asp	Leu	Phe	Ser	Val 15	Ala
	Ala	Cys	His	Arg 20	Phe											
60																

	(2) INFORMATION FOR SEQ ID NO: 289:
5	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 21 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 289:
10	Tyr Met Gly Lys Gly Ser Met Thr Gly Leu Ala Leu Lys His Met Phe 1 5 10 15
15	Glu Arg Ser Phe Thr 20
20	(2) INFORMATION FOR SEQ ID NO: 290: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 27 amino acids (B) TYPE: amino acid
25	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 290: Val Thr Gly Ile Ile Asp Ser Leu/Thr Ile Ser Pro Lys Ala Ala Arg 1 5 10 15
30	Val Gly Leu Leu Gln Tyr Ser Thr Gln Val His 20 25
35	(2) INFORMATION FOR SEQ ID NO: 291: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 24 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear
40	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 291: Thr Glu Phe Thr Leu Arg Asn Phe Asn Ser Ala Lys Asp Met Lys Lys 1 5 10 15
45	Ala Val Ala His Met Lys Tyr Met 20
50	(2) INFORMATION FOR SEQ ID NO: 292: (i) SEQUENCE CHARACTERISTICS:
55	(A) LENGTH: 27 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 292:
60	Gly Lys Gly Ser Met Thr Gly Leu Ala Leu Lys His Met Phe Glu Arg 1 5 10 15

Ser Phe Thr Gln Gly Glu Gly Ala Arg Pro Phe 5 (2) INFORMATION FOR SEQ ID NO: 293: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 44 amino acids 10 (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 293: Ser Thr Arg Val Pro Arg Ala Ala Ile Val Phe Thr Asp Gly Arg Ala 15 Gln Asp Asp Val Ser Glu Trp Ala Ser Lys Ala Lys Ala Asn Gly Ile 20 Thr Met Tyr Ala Val Gly Val Gly Lys Ala Ile Glu 35 40 25 (2) INFORMATION FOR SEQ ID NO: 294: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 42 amino acids (B) TYPE: amino acid 30 (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 294: Glu Glu Leu Gln Glu Ile Ala Ser Glu Pro Thr Asn Lys His Leu Phe 35 Tyr Ala Glu Asp Phe Ser Thr Met Asp Glu Ile Ser Glu Lys Leu Lys Lys Gly Ile Cys Glu Ala Leu Glu Asp Ser 40 35 (2) INFORMATION FOR SEQ ID NO: 295: 45 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 11 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear 50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 295: Thr Gln Arg Leu Glu Glu Met Thr Gln Arg Met 5 55

(2) INFORMATION FOR SEQ ID NO: 296:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 amino acids

	(D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 296:	
5	Pro Gln Gly Cys Pro Glu Gln Pro Leu His 1 5 10	
10	(2) INFORMATION FOR SEQ ID NO: 297:	
15	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 33 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 297:	
20	arg Cys Lys Lys Cys Thr Glu Gly Pro Ile Asp Leu Val Phe Val II 1 5 10 15	:e
	Asp Gly Ser Lys Ser Leu Gly Glu Glu Asn Phe Glu Val Val Lys Gl 20 25 30	.n
25	Phe '	
30	2) INFORMATION FOR SEQ ID NO: 298:	
	(i) SPONENCE CHARACTERICE.	
35	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 60 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 298:	
	(A) LENGTH: 60 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear	a
	(A) LENGTH: 60 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 298: Met Ala Ala Leu Leu Leu Arg His Val Gly Arg His Cys Leu Arg Al	
35	(A) LENGTH: 60 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 298: Let Ala Ala Leu Leu Leu Arg His Val Gly Arg His Cys Leu Arg Al 1 5 10 15 Lis Phe Ser Pro Gln Leu Cys Ile Arg Asn Ala Val Pro Leu Gly Th 20 25 30 Arr Ala Lys Glu Glu Met Glu Arg Phe Trp Asn Lys Asn Ile Gly Sei 35 40 45	r
35	(A) LENGTH: 60 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 298: Let Ala Ala Leu Leu Leu Arg His Val Gly Arg His Cys Leu Arg Al 1 5 10 15 Lis Phe Ser Pro Gln Leu Cys Ile Arg Asn Ala Val Pro Leu Gly Th 20 25 30 Ar Ala Lys Glu Glu Met Glu Arg Phe Trp Asn Lys Asn Ile Gly Seg	r
35	(A) LENGTH: 60 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 298: Let Ala Ala Leu Leu Leu Arg His Val Gly Arg His Cys Leu Arg Al 1 5 10 15 Lis Phe Ser Pro Gln Leu Cys Ile Arg Asn Ala Val Pro Leu Gly Th 20 25 30 Arr Ala Lys Glu Glu Met Glu Arg Phe Trp Asn Lys Asn Ile Gly Sex 35 40 45 Sn Arg Pro Leu Ser Pro His Ile Thr Ile Tyr Ser	r
35 40 45	(A) LENGTH: 60 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 298: Let Ala Ala Leu Leu Leu Arg His Val Gly Arg His Cys Leu Arg Al 1 5 10 15 Lis Phe Ser Pro Gln Leu Cys Ile Arg Asn Ala Val Pro Leu Gly Th 20 25 30 Arr Ala Lys Glu Glu Met Glu Arg Phe Trp Asn Lys Asn Ile Gly Sex 35 40 45 sn Arg Pro Leu Ser Pro His Ile Thr Ile Tyr Ser 50 55 60	r

(B) TYPE: amino acid

```
Trp Asp Leu Gly Lys Gly Leu Lys Ile Pro Gln Leu Tyr Gln Ser Gly
                                      25
                  20
 5
      (2) INFORMATION FOR SEQ ID NO: 300:
10
             (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 17 amino acids
                    (B) TYPE: amino acid
15
                   (D) TOPOLOGY: linear
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 300:
     Met Ala Ala Leu Leu Leu Arg His Val Gly Arg His Cys Leu Arg Ala
                                           10
20
     His
25
      (2) INFORMATION FOR SEQ ID NO: 301:
             (i) SEQUENCE CHARACTERISTICS:
                     (A) LENGTH: 18 amino acids
30
                     (B) TYPE: amino acid
                     (D) TOPOLOGY: linear
              (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 301:
      Val Lys Ser Leu Cys Leu Gly Pro Ala Leu Ile His Thr Ala Lys Phe
35
                        5
      Ala Leu
40
       (2) INFORMATION FOR SEQ ID NO: 302:
              (i) SEQUENCE CHARACTERISTICS:
45
                     (A) LENGTH: 23 amino acids
                     (B) TYPE: amino acid
                     (D) TOPOLOGY: linear
              (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 302:
      Val Phe Pro Leu Met Tyr His Thr Trp Asn Gly Ile Arg His Leu Met
 50
                         5
        1
       Trp Asp Leu Gly Lys Gly Leu
                    20
 55
       (2) INFORMATION FOR SEQ ID NO: 303:
              (i) SEQUENCE CHARACTERISTICS:
 60
```

```
(A) LENGTH: 22 amino acids
                      (B) TYPE: amino acid
                      (D) TOPOLOGY: linear
               (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 303:
  5
       Arg Val Trp Asp Val Arg Pro Phe Ala Pro Lys Glu Arg Cys Val Lys
                                  10
       Ile Phe Gln Gly Asn Val
 10
                    20
       (2) INFORMATION FOR SEQ ID NO: 304:
. 15
               (i) SEQUENCE CHARACTERISTICS:
                     (A) LENGTH: 30 amino acids
                      (B) TYPE: amino acid
                      (D) TOPOLOGY: linear
 20
              (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 304:
       His Asn Phe Glu Lys Asn Leu Leu Arg Cys Ser Trp Ser Pro Asp Gly
                                            10
 25
       Ser Lys Ile Ala Ala Gly Ser Ala Asp Arg Phe Val Tyr Val
                   20
 30
       (2) INFORMATION FOR SEQ ID NO: 305:
              (i) SEQUENCE CHARACTERISTICS:
                     (A) LENGTH: 30 amino acids
                     (B) TYPE: amino acid
 35
                     (D) TOPOLOGY: linear
              (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 305:
       Trp Asp Thr Thr Ser Arg Arg Ile Leu Tyr Lys Leu Pro Gly His Ala
                                           10
40
      Gly Ser Ile Asn Glu Val Ala Phe His Pro Asp Glu Pro Ile
                   20
45
       (2) INFORMATION FOR SEQ ID NO: 306:
              (i) SEQUENCE CHARACTERISTICS:
                     (A) LENGTH: 20 amino acids
50
                     (B) TYPE: amino acid
                     (D) TOPOLOGY: linear
              (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 306:
      Val Arg Gly Arg Thr Val Leu Arg Pro Gly Leu Asp Ala Glu Pro Glu
55
        1
                        5
      Leu Ser Pro Glu
60
```

```
(2) INFORMATION FOR SEQ ID NO: 307:
             (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 19 amino acids
5
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 307:
     Glu Gln Arg Val Leu Glu Arg Lys Leu Lys Lys Glu Arg Lys Lys Glu
10
                                           10
                  5
      Glu Arg Gln
15
      (2) INFORMATION FOR SEQ ID NO: 308:
             (i) SEQUENCE CHARACTERISTICS:
20
                    (A) LENGTH: 13 amino acids
                     (B) TYPE: amino acid
                     (D) TOPOLOGY: linear
              (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 308:
25
      Arg Leu Arg Glu Ala Gly Leu Val Ala Gln His Pro Pro
                        5
30
       (2) INFORMATION FOR SEQ ID NO: 309:
              (i) SEQUENCE CHARACTERISTICS:
                     (A) LENGTH: 17 amino acids
35
                     (B) TYPE: amino acid
                     (D) TOPOLOGY: linear
              (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 309:
       Gly Arg Ile Pro Ala Pro Ala Pro Ser Val Pro Ala Gly Pro Asp Ser
40
                                            10
                         5
       Arg
 45
       (2) INFORMATION FOR SEQ ID NO: 310:
              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 42 amino acids
 50
                      (B) TYPE: amino acid
                      (D) TOPOLOGY: linear
              (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 310:
       Thr Gly Cys Val Leu Val Leu Ser Arg Asn Phe Val Gln Tyr Ala Cys
 55
       Phe Gly Leu Phe Gly Ile Ile Ala Leu Gln Thr Ile Ala Tyr Ser Ile
 60
```

```
Leu Trp Asp Leu Lys Phe Leu Met Arg Asn
35 40
```

5 (2) INFORMATION FOR SEQ ID NO: 311:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 55 amino acids

10 (B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 311:

Ser Arg Ser Glu Gly Lys Ser Met Phe Ala Gly Val Pro Thr Met Arg

15 1 5 10 15

Glu Ser Ser Pro Lys Gln Tyr Met Gln Leu Gly Gly Arg Val Leu Leu 20 25 30

20 Val Leu Met Phe Met Thr Leu Leu His Phe Asp Ala Ser Phe Phe Ser 35 40 45

Ile Val Gln Asn Ile Val Gly
50 55

25

35

(2) INFORMATION FOR SEQ ID NO: 312:

30 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 60 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 312:

Gly Thr Ala Glu Asp Phe Ala Asp Gln Phe Leu Arg Val Thr Lys Gln

1 10 15

Tyr Leu Pro His Val Ala Arg Leu Cys Leu Ile Ser Thr Phe Leu Glu
20 25 30

Asp Gly Ile Arg Met Trp Phe Gln Trp Ser Glu Gln Arg Asp Tyr Ile 35 40 45

45 Asp Thr Thr Trp Asn Cys Gly Tyr Leu Leu Ala Ser 50 55 60

50 (2) INFORMATION FOR SEQ ID NO: 313:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 17 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 313:

Ala Ser Phe Leu Leu Ser Arg Thr Ser Trp Gly Thr Ala Leu Met Ile

1 5 10 15

55

Leu

```
5
      (2) INFORMATION FOR SEQ ID NO: 314:
             (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 8 amino acids
10
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 314:
      Leu Met Arg Asn Glu Ser Arg Ser
15
                       5
      . 1
      (2) INFORMATION FOR SEQ ID NO: 315:
20
             (i) SEQUENCE CHARACTERISTICS:
                     (A) LENGTH: 13 amino acids
                     (B) TYPE: amino acid
                     (D) TOPOLOGY: linear
25
              (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 315:
      Ala Ser Phe Leu Leu Ser Arg Thr Ser Trp Gly Thr Ala
                        5
30
       (2) INFORMATION FOR SEQ ID NO: 316:
              (i) SEQUENCE CHARACTERISTICS:
35
                     (A) LENGTH: 20 amino acids
                     (B) TYPE: amino acid
                     (D) TOPOLOGY: linear
              (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 316:
      Phe Ile Ser Phe Ala Asn Ser Arg Ser Ser Glu Asp Thr Lys Gln Met
40
                                            10
                        5
        1
      Met Ser Ser Phe
 45
       (2) INFORMATION FOR SEQ ID NO: 317:
              (i) SEQUENCE CHARACTERISTICS:
 50
                      (A) LENGTH: 27 amino acids
                      (B) TYPE: amino acid
                      (D) TOPOLOGY: linear
              (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 317:
 55
       Asp Pro Arg Arg Pro Asn Lys Val Leu Arg Tyr Lys Pro Pro Pro Ser
                         5
         1
       Glu Cys Asn Pro Ala Leu Asp Asp Pro Thr Pro
 60
                                        25
```

5	(2) INFORMATION FOR SEQ ID NO: 318:
10	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 30 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 318:
	Asp Tyr Met Asn Leu Leu Gly Met Ile Phe Ser Met Cys Gly Leu Met
	1 5 10 15
15	Leu Lys Leu Lys Trp Cys Ala Trp Val Ala Val Tyr Cys Ser 20 25 , 30
-	1
20	(2) INFORMATION FOR SEQ ID NO: 319:
25	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 22 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 319:
	Met Leu Ser Ile Ser Ala Val Val Met Ser Tyr Leu Gln Asn Pro Gln 1 5 10 15
30	Pro Met Thr Pro Pro Trp
	20
35	(2) INFORMATION FOR SEQ ID NO: 320:
40	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 52 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 320:
45	Ala Ala Gly Asp Gly Asp Val Lys Leu Gly Thr Leu Gly Ser Gly Ser 1 5 10 15
	Glu Ser Ser Asn Asp Gly Gly Ser Glu Ser Pro Gly Asp Ala Gly Ala 20 25 30
50	Ala Ala Xaa Gly Gly Gly Trp Ala Ala Ala Ala Leu Ala Leu Leu Thr 35 40 45
55	Gly Gly Glu 50
	(2) INFORMATION FOR SEQ ID NO: 321:
60	(i) SEQUENCE CHARACTERISTICS:

				(1	B) T	PE:	ami	77 ar no a	cid	aci	đs		•		,	
			(xi)	-	-			line PTIO		EQ II	ON C	: 321	, ' L:		•	
5	Ala 1	Ala	Asp	Asn	Tyr 5	Gly	Ile	Pro	Arg	Ala ,10	Cys	Arg	Asn	Ser	Ala 15	Arg
10	Ser	Tyr	Gly	Ala 20		Trp	Leu	Leu	Leu 25	Xaa	Pro	Ala	Gly	Ser 30	Ser ,	Arg
	Val	Glu	Pro 35	Thr	Gln	Asp	Ile	Ser 40	Ile	Ser	Åsp	Gln ,	Leu 45	Gly	Gly	Gln
15	Asp	Val 50	Pro	Val	Phe	Arg	Asn 55	Leu	Ser	Leu	Leu	Val 60	Val	Gly	Val	Gly
20	Ala 65	Val	Phe	Ser	Leú	Leu 70	Phe	His	Leu	Gly	Thr 75	Arg	Glu	Arg	Arg	Arg 80
20	Pro	His	Ala	Xaa	Glu 85	Pro	Gly	Glu	His	Thr 90	Pro	Leu	Leu	Ala	Pro 95	Ala (
25	Thr	Ala	Gln	Pro 100	Leu	Leu	Leu	Trp	Lys 105	His	Trp	Leu	Arg '	Glu 110	Xaa	Ala
	Phe	Tyr	Gln 115	Val	Ġly	Ile	Leu	Туr 120	Met	Thr	Thr	Arg	Leu 125	Ile	Val	Asn
30	Leu	Ser 130		Thr	Tyr	Met	Ala 135	Met	Tyr	Leu	Thr	Tyr 140	Ser	Leu	His	Leu
35	Pro 145	-	Lys	Phe	Ile	Ala 150		Ile	Pro	Leu	Val 155	Met	Tyr	Leu	Ser	Gly 160
<i>33</i>	Phe	Leu	Ser	Ser	Phe 165	Leu	Met	Lys	Pro	Ile 170		Lys	Cys	Ile	Gly 175	Arg
40	Asn															
	(2)	INF	ORMA	TION	FOR	SEQ	ID	NO:	322:							
45			(i)	((A) I (B) I	ENGT YPE:	TH: 2	ERIS 243 a ino a	mino acid		ids					
			(xi)					: lir [PTIC		EQ I	D NC): 32	2:			
50	Arg 1		e Thr	Asp	Asn 5		Glu	Gly	Lys	Trp		Gly	Arg	Thr	Ala 15	
55	Gly	Ser	туг	Gly		Ile	Lys	Thr	Thr 25		Val	Glu	Ile	Хаа 30		Asr
	Ser	Leu	Lys 35		Lys	Lys	Asp	Ser 40		r Gly	/ Ala	Pro	Ser 45		Pro	Ile
60	Glu	. Ast	asa c	Gln	Glu	. Va]	Тух	. Asp) Asp	val	L Ala	Glu	Gln	Asp	Asp	Ile

DEICHACIDE JAIA - NOADTOOA4 I .

		50)		•		55	;				60				
5	Ser 65	Ser	His	Ser	Gln	Ser 70		' Ser	Gly	Gly	Ile 75		Pro	Pro	Pro	Pr 8
	Asp	Asp) Asp	Ile	Tyr 85		Gly	Ile	Glu	Glu 90		Asp	Ala	Asp ,	Asp 95	
10	Phe	Pro	Ala	Pro 100	Pro	Lys	Gln	Leu	Asp 105	Met	Gly	Asp	Glu	Val		As
	Asp	Val	Asp 115	Thr	Ser	Asp	Phe	Pro 120	Val	Ser	Ser	Ala	Glu 125		Ser	Gli
15	Gly	Thr 130	Asn	Val	Gly	Lys	Ala 135	Lys	Thr	Glu	Glu	Lys 140	, Asp	Leu	Lys	Ly
20	Leu 145	Lys	Lys	Gln	Xaa	Lys 150		Xaa	Lys	Asp	Phe 155	Arg	Lys	Lys	Phe	Ly:
	Tyr	Asp	Gly	Glu	Ile 165	Arg	Val	Leu	Tyr	Ser 170	Thr	Lys	Val	Thr	Thr 175	Sei
25	Ile	Thr	Ser	Lys 180	Lys	Trp	Gly	Thr	Arg 185	Asp	Leu	Gln	Val	Lys 190	Pro	Gly
	Glu	Ser	Leu 195	Glu	Val	Ile	Gln	Thr 200	Thr	Asp	Asp	Thr	Lys 205	Val	Leu	Суя
30		210			Gly		215					220				
35	Asp 225	Asn	Asp	Gly	Glu	Ile 230	Tyr	Asp	Asp	Ile	Ala 235	Asp	Gly	Cys	Ile	Тут 240
	Asp		-			•										
40	(2)			SEQUI	FOR ENCE A) Li	CHAI	RACTI	ERIS	rics		3 _					
45			(xi)	(1 (1	B) IN D) IX JENCE	PE:	ami OGY:	no ao line	cid ear			323	3 :			
50	Ser 1	Met	Ser	Ala	Leu 5	Thr	Arg	Leu	Ala	Ser 10	Phe	Ala	Arg	Val	Gly 15	Gly
	Arg	Leu	Phe	Arg 20	Ser	Gly	Cys	Ala	Arg 25	Thr	Ala	Gly	Asp	Gly -30	Gly	Val
55	Arg	His	Ala 35	Gly	Gly	Gly	Val	His 40	Ile	Glu	Pro	Arg	Tyr 45	Arg	Gln	Phe
	Pro	Gln 50	Leu	Thr	Arg	Ser	Gln 55	Val	Phe	G1n	Ser	Glu 60	Phe	Phe	Ser	Gly
60	Leu 1	Met	Trp	Phe	Trp	Ile	Leu	Trp .	Arg	Phe	Trp	His .	Asp	Ser	Glu	Glu

369

65 70 75 80

Val Leu Gly His Phe Pro Tyr Pro Asp Pro Ser Gln Trp Thr Asp Glu 85 90 95'

Glu Leu Gly Ile Pro Pro Asp Asp Glu Asp 100 105

		· · · · · · · · · · · · · · · · · · ·		
Applicant's or agent's tile reference number	2004PCT	International application	Unassigned	
		J		

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

		line N/A
B. IDENTIFIC	CATION OF DEPOSIT	Further deposits are identified on an additional sheet
Name of deposit		Culture Collection
Address of depo	sitary institution (including postal co	ode and country)
10801 Univers Manassas, Virg United States of	zinia 20110-2209	· · · · · · · · · · · · · · · · · · ·
Date of deposit	March 7, 1997	Accession Number , 97923
C. ADDITIO	NAL INDICATIONS (leave blank	k if not applicable) This information is continued on an additional sheet
		1
		1.
		1
D DESIGNA	TED STATES FOR WHICH II	NDICATIONS ARE MADE (if the indications are not for all designated States)
DI DEGIGITA	TED STATES FOR WINCH II	NDICATIONS ARE WADE (if the indications are not for all designated States)
D 050 4 D 4 T		
E. SEPARAI	E FURNISHING OF INDICAT	CIONS (lame blank if not applicable)
	isted below will be submitted to the	International Bureau later (specify the general nature of the indications, e.g., "Accession
The indications l	isted below will be submitted to the	
The indications l	isted below will be submitted to the	
The indications l	isted below will be submitted to the	
The indications l	isted below will be submitted to the	
The indications l	isted below will be submitted to the	International Bureau later (specify the general nature of the indications, e.g., "Accessed
The indications l	isted below will be submitted to the	
The indications I Number of Deposit	isted below will be submitted to the	International Bureau later (specify the general nature of the indications, e.g., "Accessing the indications of the indication of the
The indications I Number of Deposit	isted below will be submitted to the ") For receiving Office use only was received with the international appli	For International Bureau use only This sheet was received by the International Bureau on:
The indications I Number of Deposit	isted below will be submitted to the ") For receiving Office use only was received with the international appli	International Bureau later (specify the general nature of the indications, e.g., "Accessing the indications of the indication of the

international application Unassigned			<u> </u>	
reference number PCT/US 3B	Applicant's or agent's tile reference number	Z004PCT	International application.	

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

			
	ons made below relate to the mi	croorganism referm line N/A	· · · · · · · · · · · · · · · · · · ·
	CATION OF DEPOSIT		
B. IDENTIFIC	CATION OF DEPOSIT	ŀ	Further deposits are identified on an additional sheet
Name of deposit	ary institution _! American	Гуре Culture Col	lection
Address of depo	sitary institution (including pos	tal code and count	(بر
10801 Univers Manassas. Virg United States of	ginia 20110-2209		i k
	1	•	
Date of deposit	May 22, 1997		Accession Number 209071
C. ADDITIO	NAL INDICATIONS (leave	blank if not applicab	ble) This information is continued on an additional sheet
		,	1
			ł.
D. DESIGNA	TED STATES FOR WHIC	CH INDICATIO	NS ARE MADE (If the indications are not for all designated States
E SEPARAT	E FURNISHING OF IND	CATIONS (legre	h lank if not applicable
The indications	listed below will be submitted t		Bureau later (specify the general nature of the indications, e.g., "Access
Number of Depos	ir)		
•			
:	٠		
<u> </u>			
	For receiving Office use o	nly	For International Bureau use only
This sheet	was received with the internations	d application	This sheet was received by the International Bureau on:
Authorized office	T.		Authorized officer
Vina	inia Ilila	.,	

<u> </u>		372
Applicant's or agent's file reference number	Z004PCT	International application Unassigned

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

A. The indication on page	ons made below relate to the microorganism ref 73 . line N	erred to in the description
R. IDENTIFI	CATION OF DEPOSIT	
Name of deposit		Further deposits are identified on an additional sheet
10801 Universi	inia 20110-2209	intry)
Date of deposit	February 25, 1998	Accession Number 209641
C. ADDITIO	NAL INDICATIONS (leave blank if not applic	cable) This information is continued on an additional sheet
D. DESIGNAT	TED STATES FOR WHICH INDICATE	ONS ARE MADE (if the indications are not for all designated States)
E. SEPARATI	E FURNISHING OF INDICATIONS (lea	Mark Company
The indications li Number of Deposit	isted below will be submitted to the Internations	we ciank if not applicable) al Burcau later (specify the general nature of the indications, e.g., "Accessive and the indications of the indications.
Authorized officer	For receiving Office use only was received with the international application	This sheet was received by the International Bureau on: Authorized officer

· · · · · · · · · · · · · · · · · · ·		
Applicant's or agent's file reference number	Z004PCT	International application Unassigned

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

	s made below relate to the microor	ganism referred to in the description . line N/A
. IDENTIFIC	ATION OF DEPOSIT	Further deposits are identified on an additional sheet
Name of deposita	y institution American Type	Culture Collection
Address of depos	itary institution (including postal co	ode and country)
10801 Universi	ty Boulevard inia 20110-2209	
	1	
Date of deposit	July 24, 1997	Accession Number 209179
C. ADDITION	NAL INDICATIONS (leave blan	k if not applicable) This information is continued on an additional sheet
· · · · · · · · · · · · · · · · · · ·		11
D. DESIGNA	CED STATES FOR WHICH I	INDICATIONS ARE MADE (if the indications are not for all designated States
- CED. D. E		TIONS 4 Habitan Habia
		TIONS (leave blank if not applicable)
The indications Number of Deposi		e International Bureau later (specify the general nature of the indications. e.g "Acces
Number of Deposi	, ,	
		For International Bureau use only
	For receiving Office use only	
This sheet	For receiving Office use only was received with the international app	plication This sheet was received by the International Bureau on:
This sheet	•	plication This sheet was received by the International Bureau on:
This sheet	was received with the international app	This sheet was received by the International Bureau on: . Authorized officer
Authorized office	was received with the international app	

BNSDOCID: <WO__9842738A1_I_> 134 (July 1992)

3/	4
Applicant's or agent's tile Z004PCT	
reference number	International application . Unassigned
	Omassigned

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

· ·	(PCT Rule 13bis)	
A. The indications made below relate to the microorganism referred to in the description on page 77 . line N/A		
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet	
Name of depositary institution American Type Cultur		
Address of depositary institution (including postal code and 10801 University Boulevard Manassas. Virginia 20110-2209 United States of America	(country)	
Date of deposit March 7, 1997	Accession Number 97924	
C. ADDITIONAL INDICATIONS (leave blank if not ap	oplicable) This information is continued on an additional sheet	
l		
D. DESIGNATED STATES FOR WHICH INDICA	TIONS ARE MADE (if the indications are not for all designated States)	
SEPARATE FURNISHING OF INDICATIONS	7 N	
he indications listed below will be submitted to the International (umber of Deposit")	onal Burcau later (specify the general nature of the indications. e.g "Accession	
For receiving Off		
For receiving Office use only	This sheet was received by the International Bureau on:	
Verginia L lely	Authorized officer	

	_ ,			
Applicant's or agent's file reference number	Z004PCT	International application	Unassigned	

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page 80 . line N/A			
on page 8		· inic i Tork	
B. IDENTIFIC	ATION OF DEPOSIT	· · · · · · · · · · · · · · · · · · ·	Further deposits are identified on an additional sheet
Name of deposita		e Culture Collec	ction
Address of depos	itary institution (including postal c	ode and country)	1
10801 Universi Manassas, Virg United States o	inia 20110-2209		
		1	1
Date of deposit	March 13, 1997	A	Accession Number 97958
C. ADDITION	NAL INDICATIONS (leave blan	nk if not applicable)	This information is continued on an additional sheet
		1	
ļ			
D. DESIGNA	TED STATES FOR WHICH	INDICATIONS	S ARE MADE (if the indications are not for all designated States)
E. SEPARAT	E FURNISHING OF INDICA	TIONS (leave b	olank if not applicable)
The indications Number of Deposi		e International Bo	ureau later (specify the general nature of the indications. e.g "Accession
	·		
	For receiving Office use only		For International Bureau use only
This sheet	was received with the international ap	plication	This sheet was received by the International Bureau on:
Authorized office			Authorized officer
Vua	mia L bely	,	

BNSDOCID: <WO__9842738A1_I_> 1/134 (July 1992)

WO 98/42738

Applicant's or agent's file reference number	Z004PCT	International application Unassigned

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

A. The indications made below relate to the microorganism referred to in the description on page 80 . line N/A		
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet	
Name of depositary institution American Type Culture Co		
Address of depositary institution (including postal code and coun	urv)	
10801 University Boulevard Manassas. Virginia 20110-2209 United States of America	· · · · · · · · · · · · · · · · · · ·	
Date of deposit May 22, 1997	Accession Number 209072	
C. ADDITIONAL INDICATIONS (leave blank if not applica-	ble) This information is continued on an additional sheet	
·	! !	
	i	
D. DESIGNATED STATES FOR WHICH INDICATIO	NS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave	blank if not applicable)	
The indications listed below will be submitted to the International Number of Deposit")	Bureau later (specify the general nature of the indications, e.g., "Accession	
For receiving Office use only	For International Bureau use only	
This sheet was received with the international application Authorized officer	This sheet was received by the International Bureau on:	
Virginia L Illey	Authorized officer	

-	-	-
4		

Applicant's or agent's file Z004PCT reference number	International application Unassigned

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page 80 line N/A		
B. IDENTIFICATION OF DEPOSIT	1	, Further deposits are identified on an additional sheet
Name of depositary institution America	ın Type Culture Collec	ction
Address of depositary institution (including	postal code and country)
10801 University Boulevard Manassas, Virginia 20110-2209 United States of America		
Date of deposit September 4, 1997	1	Accession Number 209235
C. ADDITIONAL INDICATIONS (le	ave blank if not applicable,	This information is continued on an additional sheet
		1
	•	,
		1
D DESIGNATED STATES FOR WE	HICH INDICATION	S ARE MADE (if the indications are not for all designated States)
D. DESIGNATED STATES FOR WI		O TIME WILL SHE SHEET SH
E. SEPARATE FURNISHING OF IT		
The indications listed below will be submitt Number of Deposit")	ed to the International B	ureau later (specify the general nature of the indications. e.g., "Accession
		·
For receiving Office us	se only	For International Bureau use only
This sheet was received with the internal	tional application	This sheet was received by the International Bureau on:
•		
Authorized officer		Authorized officer

BNSDOCID: <WO__9842738A1_I_>

	37	78 .		
Applicant's or agent's file reference number	Z004PCT	International application	Unassigned	·
			<u>'''도 화</u> 공:	<u></u>

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

A. The indications made below relate to the microorganism referred to in the description on page 84 . line N/A		
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet	
Name of depositary institution American Type Culture Col		
Address of depositary institution (including postal code and count 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America		
Date of deposit August 28, 1997	Accession Number 209226	
C. ADDITIONAL INDICATIONS (leave blank if not applicab	This information is continued on an additional sheet	
I		
D. DESIGNATED STATES FOR WHICH INDICATION	NS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave	blank if not applicable)	
The indications listed below will be submitted to the International E Number of Deposit*)	Sureau later (specify the general nature of the indications, e.g., "Accession	
For receiving Office use only This sheet was received with the international application Authorized officer Luguna Llly	For International Bureau use only This sheet was received by the International Bureau on: Authorized officer	

Applicant's or agent's tile	Z004PCT	International application Unassigned
reference number		PCT/US 98/

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

on page	84	microorganism refer	red to in the description
3. IDENTIFIC	CATION OF DEPOSIT		Further deposits are identified on an additional sheet
Name of deposite	ary institution America	un Type Culture Co	llection
10801 Univers	ginia 20110-2209	postal code and coun	try)
	,	1	1
Date of deposit	March 13, 1997		Accession Number 97957
C. ADDITIO	NAL INDICATIONS (le	ave blank if not applica	ble) This information is continued on an additional sheet
		ı	
			,
D DECICAL			
D. DESIGNA	TED STATES FOR WE	HICH INDICATIO	ONS ARE MADE (if the indications are not for all designated States
D. DESIGNA	TED STATES FOR WE	HICH INDICATIO	ONS ARE MADE (if the indications are not for all designated States
D. DESIGNA	TED STATES FOR WE	ich indica i c	ONS ARE MADE (if the indications are not for all designated States
. DESIGNA	TED STATES FOR WE	HICH INDICATIO	ONS ARE MADE (if the indications are not for all designated States
D. DESIGNA	TED STATES FOR WE	HICH INDICATIO	ONS ARE MADE (if the indications are not for all designated States
E. SEPARAT	TE FURNISHING OF IN	NDICATIONS (lea	ve blank if not applicable)
E. SEPARAT	TE FURNISHING OF IN	NDICATIONS (lea	
E. SEPARAT	TE FURNISHING OF IN	NDICATIONS (lea	ve blank if not applicable)
E. SEPARAT	TE FURNISHING OF IN	NDICATIONS (lea	ve blank if not applicable)
E. SEPARAT	TE FURNISHING OF IN	NDICATIONS (lea	ve blank if not applicable)
E. SEPARAT	TE FURNISHING OF IN	NDICATIONS (lea	ve blank if not applicable)
E. SEPARAT	TE FURNISHING OF IN	NDICATIONS (lear	ve blank if not applicable)
E. SEPARAT The indications Number of Depos	TE FURNISHING OF IN listed below will be submitted.	NDICATIONS (learned to the International	ve blank if not applicable) Il Bureau later (specify the general nature of the indications. e.g., "Acces
E. SEPARAT The indications Number of Depos	TE FURNISHING OF IN listed below will be submitt it") For receiving Office us at was received with the internat	NDICATIONS (learned to the International	we blank if not applicable) Il Bureau later (specify the general nature of the indications, e.g., "Access For International Bureau use only

Applicant's or agent's file Z004PCT reference number	International application Unassigned
	PCT/US 98:

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

1	(PCI Rule 1366)
A. The indications made below relate to the microorganism on page 84 , line	m referred to in the description N/A
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet
Name of depositary institution American Type Cultu	
Address of depositary institution (including postal code and 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	d country)
Date of deposit May 22, 1997	Accession Number 209073
C. ADDITIONAL INDICATIONS (leave blank if not a	applicable) This information is continued on an additional sheet
	1
D. DESIGNATED STATES FOR WHICH INDICA	ATIONS ARE MADE (if the indications are not for all designated States)
E. SEPARATE FURNISHING OF INDICATIONS The indications listed below will be submitted to the Internal	o (leave blank if not applicable) tional Bureau later (specify the general nature of the indications, e.g., "Accession
	·
For receiving Office use only	For International Bureau use only
This sheet was received with the international application uthorized officer	This sheet was received by the International Bureau on:
Vuginia Llily	Authorized officer

What Is Claimed Is:

5

10

15

20

25

- 1. An isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 95% identical to a sequence selected from the group consisting of:
- (a) a polynucleotide fragment of SEQ ID NO:X or a polynucleotide fragment of the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
- (b) a polynucleotide encoding a polypeptide fragment of SEQ ID NO:Y or a polypeptide fragment encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
- (c) a polynucleotide encoding a polypeptide domain of SEQ ID NO:Y or a polypeptide domain encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
- (d) a polynucleotide encoding a polypeptide epitope of SEQ ID NO:Y or a polypeptide epitope encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
 - (e) a polynucleotide encoding a polypeptide of SEQ ID NO:Y or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X, having biological activity;
 - (f) a polynucleotide which is a variant of SEQ ID NO:X;
 - (g) a polynucleotide which is an allelic variant of SEQ ID NO:X;
 - (h) a polynucleotide which encodes a species homologue of the SEQ ID NO:Y;
 - (i) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h), wherein said polynucleotide does not hybridize under stringent conditions to a nucleic acid molecule having a nucleotide sequence of only A residues or of only T residues.
 - 2. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding a secreted protein.
 - 3. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding the sequence identified as SEQ ID NO:Y or the polypeptide encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.

4. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises the entire nucleotide sequence of SEQ ID NO:X or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.

5

5. The isolated nucleic acid molecule of claim 2, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.

10

- 6. The isolated nucleic acid molecule of claim 3, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.
- 7. A recombinant vector comprising the isolated nucleic acid molecule of
 - 8. A method of making a recombinant host cell comprising the isolated nucleic acid molecule of claim 1.

20

- A recombinant host cell produced by the method of claim 8.
- 10. The recombinant host cell of claim 9 comprising vector sequences.
- 11. An isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence selected from the group consisting of:
 - (a) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
 - (b) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z, having biological activity;

- (c) a polypeptide domain of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
- (d) a polypeptide epitope of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
- (e) a secreted form of SEQ ID NO:Y or the encoded sequence included inATCC Deposit No:Z;
 - (f) a full length protein of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

- (g) a variant of SEQ ID NO:Y;
- (h) an allelic variant of SEQ ID NO:Y; or
- (i) a species homologue of the SEQ ID NO:Y.
- The isolated polypeptide of claim 11, wherein the secreted form or the full length protein comprises sequential amino acid deletions from either the C-terminus or the N-terminus.
 - 13. An isolated antibody that binds specifically to the isolated polypeptide of claim 11.
- 10

 14. A recombinant host cell that expresses the isolated polypeptide of claim
 11.
 - 15. A method of making an isolated polypeptide comprising:
- 15 (a) culturing the recombinant host cell of claim 14 under conditions such that said polypeptide is expressed; and
 - (b) recovering said polypeptide.
 - 16. The polypeptide produced by claim 15.
 - 17. A method for preventing, treating, or ameliorating a medical condition, comprising administering to a mammalian subject a therapeutically effective amount of the polypeptide of claim 11 or the polynucleotide of claim 1.
- 25 18. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:
 - (a) determining the presence or absence of a mutation in the polynucleotide of claim 1; and
- (b) diagnosing a pathological condition or a susceptibility to a pathologicalcondition based on the presence or absence of said mutation.
 - 19. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:
- (a) determining the presence or amount of expression of the polypeptide of claim 11 in a biological sample; and
 - (b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or amount of expression of the polypeptide.

- 20. A method for identifying a binding partner to the polypeptide of claim 11 comprising:
 - (a) contacting the polypeptide of claim 11 with a binding partner; and
- 5 (b) determining whether the binding partner effects an activity of the polypeptide.
 - 21. The gene corresponding to the cDNA sequence of SEQ ID NO:Y.
- 10 22. A method of identifying an activity in a biological assay, wherein the method comprises:
 - (a) expressing SEQ ID NO:X in a cell;
 - (b) isolating the supernatant;
 - (c) detecting an activity in a biological assay; and
 - (d) identifying the protein in the supernatant having the activity.
 - 23. The product produced by the method of claim 22.



PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent	Classification 6:	Ι.	(1	(1) International Publication Number: WO 98/42738
	1, C07K 14/47, 16/18, C12Q /53, 33/68, A61K 38/17	A1	(4	13) International Publication Date: 1 October 1998 (01.10.98)
(21) International Applica		S98/053 (19.03.9		60/056,370 19 August 1997 (19.08.97) US 60/060,862 2 October 1997 (02.10.97) US (71) Applicant (for all designated States except US): HUMAN
(30) Priority Data: 60/041,281 60/042,344 60/041,277 60/048,355 60/048,096 60/048,351 60/048,154 60/048,160 60/048,131 60/048,186 60/048,095 60/048,187 60/048,352 60/048,135 60/048,135 60/048,352 60/048,188 60/048,350 60/048,350 60/054,804	21 March 1997 (21.03.97) 21 March 1997 (21.03.97) 21 March 1997 (21.03.97) 21 March 1997 (21.03.97) 30 May 1997 (30.05.97) 5 August 1997 (05.08.97)		US U	GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): YOUNG, Paul [US/US]; 122 Beckwith Street, Gaithersburg, MD 20878 (US). GREENE, John, M. [US/US]; 872 Diamond Avenue, Gaithersburg, MD 20878 (US). FERRIE, Ann, M. [US/US]; 13203 L Astoria Hill Court, Germantown, MD 20874 (US). RUBEN, Steven, M. [US/US]; 18528 Heritage Hills Drive, Olney, MD 20832 (US). ROSEN, Craig, A. [US/US]; 22400 Rolling Hill Road, Laytonsville, MD 20882 (US). DUAN, Roxanne [US/US]; 4541 Fairfield Drive, Bethesda, MD 20814 (US). HU, Jing-Shan [CN/US]; 1247 Lakeside Drive #3034, Sunnyvale, CA 94086 (US). FLORENCE, Kimberly, A. [US/US]; 12805 Atlantic Avenue, Rockville, MD 20851 (US). OLSEN, Henrik, S. [DK/US]; 182 Kendrick Place #24, Gaithersburg, MD 20878 (US). EBNER, Reinhard [DE/US]; 9906 Shelburne Terrace #316, Gaithersburg, MD 20878 (US). BREWER, Laurie, A. [US/US]; 14920 M. Nebo Road, Poolesville, MD 20837 (US). MOORE, Paul, A. [GB/US]; Apartment 104, 1908 Holly Ridge Drive, McLean, VA 22102 (US). SHI, Yanggu [CN/US]; 437 West Side Drive, Gaithersburg, MD 20878 (US). LAFLEUR, David, W. [US/US]; 1615 Q Street, N.W. #807, Washington, DC 20009 (US). NI, Jian [CN/US]; 5502 Manorfield Road, Rockville, MD 20853 (US). (74) Agents: BROOKES, Anders, A. et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 10850 (US). (81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), Eu
				Published With international search report.

(54) Title: 87 HUMAN SECRETED PROTEINS

(57) Abstract

The present invention relates to 87 novel human secreted proteins and isolated nucleic acids containing the coding regions of the genes encoding such proteins. Also provided are vectors, host cells, antibodies, and recombinant methods for producing human secreted proteins. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating disorders related to these novel human secreted proteins.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM AT AU	Armenia	-		LS	Lesotho	SI	Slovenia
1		FI	Finland	LT	Lithuania	SK	Slovakia
ATT	Austria	FR	Prance	LU	Luxembourg	SN	Senegal
	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Ita ly	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
СН	Switzerland	KG	Kyrgyzstan	NO	Norway	zw	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand	2**	Zimoaowe
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Pederation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

WO 98/42738 PCT/US98/05311

87 Human Secreted Proteins

Field of the Invention

This invention relates to newly identified polynucleotides and the polypeptides encoded by these polynucleotides, uses of such polynucleotides and polypeptides, and their production.

Background of the Invention

Unlike bacterium, which exist as a single compartment surrounded by a membrane, human cells and other eucaryotes are subdivided by membranes into many functionally distinct compartments. Each membrane-bounded compartment, or organelle, contains different proteins essential for the function of the organelle. The cell uses "sorting signals," which are amino acid motifs located within the protein, to target proteins to particular cellular organelles.

One type of sorting signal, called a signal sequence, a signal peptide, or a leader sequence, directs a class of proteins to an organelle called the endoplasmic reticulum (ER). The ER separates the membrane-bounded proteins from all other types of proteins. Once localized to the ER, both groups of proteins can be further directed to another organelle called the Golgi apparatus. Here, the Golgi distributes the proteins to vesicles, including secretory vesicles, the cell membrane, lysosomes, and the other organelles.

Proteins targeted to the ER by a signal sequence can be released into the extracellular space as a secreted protein. For example, vesicles containing secreted proteins can fuse with the cell membrane and release their contents into the extracellular space - a process called exocytosis. Exocytosis can occur constitutively or after receipt of a triggering signal. In the latter case, the proteins are stored in secretory vesicles (or secretory granules) until exocytosis is triggered. Similarly, proteins residing on the cell membrane can also be secreted into the extracellular space by proteolytic cleavage of a "linker" holding the protein to the membrane.

Despite the great progress made in recent years, only a small number of genes encoding human secreted proteins have been identified. These secreted proteins include the commercially valuable human insulin, interferon, Factor VIII, human growth hormone, tissue plasminogen activator, and erythropoeitin. Thus, in light of the pervasive role of secreted proteins in human physiology, a need exists for identifying and characterizing novel human secreted proteins and the genes that encode them. This knowledge will allow one to detect, to treat, and to prevent medical disorders by using secreted proteins or the genes that encode them.

5

10

15

20

25

30

10

15

20

25

30

35

Summary of the Invention

The present invention relates to novel polynucleotides and the encoded polypeptides. Moreover, the present invention relates to vectors, host cells, antibodies, and recombinant methods for producing the polypeptides and polynucleotides. Also provided are diagnostic methods for detecting disorders related to the polypeptides, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying binding partners of the polypeptides.

Detailed Description

Definitions

The following definitions are provided to facilitate understanding of certain terms used throughout this specification.

In the present invention, "isolated" refers to material removed from its original environment (e.g., the natural environment if it is naturally occurring), and thus is altered "by the hand of man" from its natural state. For example, an isolated polynucleotide could be part of a vector or a composition of matter, or could be contained within a cell, and still be "isolated" because that vector, composition of matter, or particular cell is not the original environment of the polynucleotide.

In the present invention, a "secreted" protein refers to those proteins capable of being directed to the ER, secretory vesicles, or the extracellular space as a result of a signal sequence, as well as those proteins released into the extracellular space without necessarily containing a signal sequence. If the secreted protein is released into the extracellular space, the secreted protein can undergo extracellular processing to produce a "mature" protein. Release into the extracellular space can occur by many mechanisms, including exocytosis and proteolytic cleavage.

As used herein, a "polynucleotide" refers to a molecule having a nucleic acid sequence contained in SEQ ID NO:X or the cDNA contained within the clone deposited with the ATCC. For example, the polynucleotide can contain the nucleotide sequence of the full length cDNA sequence, including the 5' and 3' untranslated sequences, the coding region, with or without the signal sequence, the secreted protein coding region, as well as fragments, epitopes, domains, and variants of the nucleic acid sequence. Moreover, as used herein, a "polypeptide" refers to a molecule having the translated amino acid sequence generated from the polynucleotide as broadly defined.

In the present invention, the full length sequence identified as SEQ ID NO:X was often generated by overlapping sequences contained in multiple clones (contig

15

20

25

30

35

analysis). A representative clone containing all or most of the sequence for SEQ ID NO:X was deposited with the American Type Culture Collection ("ATCC"). As shown in Table 1, each clone is identified by a cDNA Clone ID (Identifier) and the ATCC Deposit Number. The ATCC is located at 10801 University Boulevard, Manassas, Virginia 20110-2209, USA. The ATCC deposit was made pursuant to the terms of the Budapest Treaty on the international recognition of the deposit of microorganisms for purposes of patent procedure.

A "polynucleotide" of the present invention also includes those polynucleotides capable of hybridizing, under stringent hybridization conditions, to sequences contained in SEQ ID NO:X, the complement thereof, or the cDNA within the clone deposited with the ATCC. "Stringent hybridization conditions" refers to an overnight incubation at 42° C in a solution comprising 50% formamide, 5x SSC (750 mM NaCl, 75 mM sodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 μg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

Also contemplated are nucleic acid molecules that hybridize to the polynucleotides of the present invention at lower stringency hybridization conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency); salt conditions, or temperature. For example, lower stringency conditions include an overnight incubation at 37°C in a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 0.2M NaH₂PO₄; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 ug/ml salmon sperm blocking DNA; followed by washes at 50°C with 1XSSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5X SSC).

Note that variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include Denhardt's reagent, BLOTTO, heparin, denatured salmon sperm DNA, and commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

Of course, a polynucleotide which hybridizes only to polyA+ sequences (such as any 3' terminal polyA+ tract of a cDNA shown in the sequence listing), or to a

10

15

20

25

30

35

complementary stretch of T (or U) residues, would not be included in the definition of "polynucleotide," since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone).

The polynucleotide of the present invention can be composed of any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of single-and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, the polynucleotide can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. A polynucleotide may also contain one or more modified bases or DNA or RNA backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications can be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

The polypeptide of the present invention can be composed of amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched, for example, as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine,

10

15

20

30

35

formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. (See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993); POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, pgs. 1-12 (1983); Seifter et al., Meth Enzymol 182:626-646 (1990); Rattan et al., Ann NY Acad Sci 663:48-62 (1992).)

"SEQ ID NO:X" refers to a polynucleotide sequence while "SEQ ID NO:Y" refers to a polypeptide sequence, both sequences identified by an integer specified in Table 1.

"A polypeptide having biological activity" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. In the case where dose dependency does exist, it need not be identical to that of the polypeptide, but rather substantially similar to the dose-dependence in a given activity as compared to the polypeptide of the present invention (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about tenfold less activity, and most preferably, not more than about three-fold less activity relative to the polypeptide of the present invention.)

25 Polynucleotides and Polypeptides of the Invention

FEATURES OF PROTEIN ENCODED BY GENE NO: 1

The translation product of this gene shares sequence homology with nucleolin, which is thought to be important in macromolecule binding, as well as some membrane proteins. Preferred polypeptide fragments comprise the amino acid sequence: DPEAADSGEPQNKRTPDLPEEEYVKEEIQENEEAVKKMLVEATREFEEVVVDES (SEQ ID NO:239); QKLKRKAEEDPEAADSGEPQNKRTPDLPEEEYVKEEIQENEE AVKKMLVEATREFEEVVVDES (SEQ ID NO:240); KAMEKSSLTQHSWQSLKDR YLKHLRGQEHKYLLGDAPVSPSSQKLKRKAEEDPEAADSGEPQNKRTPDLPEE EYVKEEIQENEEAVKKMLVEATREFEEVVVDESPPDFEIHI (SEQ ID NO:241). Also preferred are the polynucleotide fragments encoding these polypeptide fragments.

10

15

20

25

30

35

This gene maps to chromosome 16, and therefore can be used as a marker in linkage analysis for chromosome 16.

This gene is expressed primarily in brain and kidney and to a lesser extent in wide range of tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, cell-cell interaction or cell-matrix interaction. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain and kidney, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., brain and other tissue of the nervous system, and kidney, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:125 as residues: Met-1 to Trp-10.

The tissue distribution and homology to nucleolin indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment/diagnosis of diseases involving cell-cell interaction or cell-extracellular matrix interaction.

FEATURES OF PROTEIN ENCODED BY GENE NO: 2

The translation product of this gene shares sequence homology with a porcine zona pellucida protein ZPDS.1711. (See Accession No. R39356.) These two proteins have weak homology with *Drosophila* commissureless and metal homeostasis proteins which are thought to be important in controlling growth cone guidance across the CNS midline and protecting cells against reactive oxygen toxicity. thus, based on homology, it is likely that this gene also be involved in development. Preferred polypeptide fragments comprise the amino acid sequence: LPSYDEAERTKAEATIPLVPGRDEDF VGRDDFDDADQLRIGNDGIFMLTFFMAFLFNWIGFFLSFCLTTSAAGRYGAISG FGLSLIKWILIVRFSTYFPGYFDGQYWLWWVFLVLGFLLFLRGFINYAKVRKM PETFSNLPRTRVLFI (SEQ ID NO:242); and/or AGRYGAISGFGLSLIKWILIVRFS (SEQ ID NO:243). Also preferred are polynucleotide fragments encoding these polypeptide fragments. This gene maps to chromosome 5, and therefore can be used in linkage analysis as a marker for chromosome 5.

10

15

20

25

30

35

This gene is expressed primarily in kidney, adrenal gland, brain and to a lesser extent in wide range of tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, fertilization control or tissue damages by metabolites or other toxic agents. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive and urosecretion system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., kidney, adrenal gland, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to zona pellucida protein indicates that polynucleotides and polypeptides corresponding to this gene are useful for fertility control such as controceptive development. The homology with metal homeostasis and commissureless genes indicates the gene's function in spermatozoa guidance and protection. It would also be useful for the treatment/diagnosis of tissue damages caused by toxic metabolites and other agents since the gene product is also expressed in urosecretive tissues.

FEATURES OF PROTEIN ENCODED BY GENE NO: 3

This gene is expressed primarily in liver and to a lesser extent in placenta. Preferred polypeptide fragments comprise the amino acid sequence: MKHLSAWNFT KLTFLQLWEI FEGSVENCQTLTSYSKLQIKYTFSRGSTFYI (SEQ ID NO:244). Also preferred are polynucleotide fragments encoding these polypeptide fragments.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, digestive and nutrient transport/utilization disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the digestive and

10

15

20

25

30

35

circulatory system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., liver, and placenta, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in liver and placenta indicates that the protein product is either an extracellular enzyme or a molecule carrier. Therefore, polynucleotides and polypeptides corresponding to this gene are useful for diagnosis/treatment of digestive and nutrient transport/utilization disorders, including malabsorption and malnutrition.

FEATURES OF PROTEIN ENCODED BY GENE NO: 4

This gene shares homology with the sap47 gene of Drosophila melanogaster, a gene which codes for a conserved neuronal protein associated with synaptic terminals. (See Mol. Brain Res. 32:45-54 (1995); see also, Accession No. 929571.) Thus, based on homology, the gene of the present invention also should be associated with synaptic terminals. Preferred polypeptide fragments comprise the amino acid sequence: FSSDFRTSPWESRRVESKATSARCGLWGSGPRRRPASGMFRGLSSWLGLQQP VAGGGQPNGDAPPEQPSETVAESAEEELQQAGDQELLHQAKDFGNYLFNFASA ATKKITESVAETAQTIKKSVEEGKIDGIIDKTIIGDFQKEQKKFVEEQHTKKSEA AVPPWVDTNDEETIQQQILALSADKRNFLRDPPAGVQFNFDFDQMYPVALVML (SEQ ID NO:245); MRFALVPKLVKEEVFWRNYFYRVSLIKQSAQLTALAAQQQA AGKGGEEQ (SEQ ID NO:246); STSPGVSEFVSDAFDACNLNQEDLRKEMEQL VLDKKQEETAVLEEDSADWEKELQQELQEYEVVTESEKRDENWDK (SEQ ID NO:247); SPWESRRVESKATSARCGLWGSGPRRRPASGMFRGLSSWLGLQQ PVAGGGQPNGDAPPEQPS (SEQ ID NO:248); PVAGGGQPNGDAPPEQPSETV ESAEELQQAGDQELLHQAKDFGNYLFNFASAATKKITESVAE (SEQ ID NO: 249); and/or FQKEQKKFVEEQHTKKSEAAVPPWVDTNDEETIQQQILALSADKR NFLRDPPAGVQFNFDFDQMYPVALVML (SEQ ID NO:250). Also preferred are polynucleotide fragments encoding these polypeptide fragments.

This gene is expressed primarily in kidney pyramids and to a lesser extent in lung and other tissues of various types. This gene fluxes calcium in human aortic smooth muscle cells, and therefore is involved in signal transduction.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are

10

15

20

25

30

35

not limited to, renal and nervous disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the kidney and/or nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., kidney, lung, brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in kidney and lung and homology with sap47 indicates that the protein product has regulatory or direct functions in molecular exchange with body fluids and nervous system signaling. Polynucleotides and polypeptides corresponding to this gene are useful for treatment of disorders in kidney and nervous system.

FEATURES OF PROTEIN ENCODED BY GENE NO: 5

The translation product of this gene shares sequence homology with the mouse Ly-9.2 antigen which is thought to be an important cell surface marker in lymphoids, myeloids and hematopoietic progenitors. (See Accession No. gil198932.) Preferred polypeptide fragments comprise the amino acid sequence: PFICVARNPVSRNFSSPI LARKLCEGAA (SEQ ID NO:251); and/or KEDPANTVYSTVEIPKKMENPHSLLT MPDTPRL (SEQ ID NO:252). Also preferred are polynucleotide fragments encoding these polypeptide fragments. Based on homology, it is likely that this gene is also a cell surface marker, involved in hematopoiesis.

This gene is expressed primarily in activated macrophages, monocytes and T-cells and to a lesser extent in spleen and bone marrow.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune and hematopoietic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hematopoietic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., blood cells, and bone marrow, and cancerous and wounded

10

15

20

25

30

35

tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:129 as residues: Lys-26 to Tyr-33, Arg-44 to Ile-49, Ser-53 to Lys-71, Lys-86 to Pro-91.

The tissue distribution and homology to Ly-9.2 surface immunoglobulin family indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis of immune and hematopoietic disorders. Polypeptides and polynucleotides corresponding to this gene are also be used as a marker for leukemia or a modulator of the functions of the cells of macrophage/monocyte or T-cell types.

FEATURES OF PROTEIN ENCODED BY GENE NO: 6

The translation product of this gene shares sequence homology with the *Drosophila* glutactin gene which is thought to be important in cell-cell interaction or cell-extracellular matrix contact.

This gene is expressed primarily in colon tissue, aorta endothelial cells and to a lesser extent in skin, breast tissue and T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of these tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, diseases of the gastrointestinal tract, vascular system or T-cell development. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of these tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the digestive system, cardiovascular system, and immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., colon, cardiovascular tissue, skin, mammary tissue, and blood cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to glutactin indicates that polynucleotides and polypeptides corresponding to this gene are useful for the development and maintenance of the integrity of the basal membrane in the gastrointestinal tract and

10

15

20

cardiovascular system. The expression in T-cells also indicate the protein may be involved in T-cell adhesion, cell-cell interaction and development.

FEATURES OF PROTEIN ENCODED BY GENE NO: 7

The translation product of this gene shares sequence homology with MURF4 protein, an ATPase homolog, which is thought to be important in ATP hydrolysis.

This gene is expressed primarily in breast tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, breast cancer and non-neoplastic breast diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of these tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the breast tissue, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., mammary tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to MURF4 gene indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of neoplastic or non-neoplastic breast diseases because ATPase like protein may be involved in changed metabolic states of the breast.

25

30

35

FEATURES OF PROTEIN ENCODED BY GENE NO: 8

This gene shares homology to the alcohol dehydrogenase gene. Preferred polypeptide fragments comprise comprise the amino acid sequence: ASAVLLDLPNSG GEAQAKKLGNNCVFAPADVTSEKDVQTALALAKGKFGRVDVAVNCAGIAVAS KTYNLKKGQTHTLEDFQRVLDVNLMGTFNVIRLVAGEMGQNEPDQGGQRGVI INTASVAAFEGQVGQAAYSASKGGIVGMTLPIARDLAPIGIRVMTIAPGLFGTPL LTSLPEKVCNFLASQVPFPSRLGDPAEYAHLVQAIIENPFLNGEVIRLDGAIRMQ P (SEQ ID NO:253); and/or SVAAFEGQVGQAAYSASKGGIVGMTLPIA (SEQ ID NO:254). Polynucleotides encoding these fragements are also encompassed by the invention. Other groups have also recently cloned this gene, recognizing its homology to alcohol dehydrogenase. (See Accession No. 1778355.) Moreover, a second group

recently cloned the mouse homologue of this gene. (See Accession No. 2078284.) They found that the mouse homologue binds to amyloid beta-peptide and mediates neurotoxicity in Alzheimer's disease, calling the protein ERAB. This gene maps to chromosome X, and therefore can be used in linkage analysis as a marker for chromosome X. Therefore, mutations in the translated product of this gene may be involved in Alzheimer's disease in humans, as well as other sex linked diseases. This gene can be used as a diagnostic marker for these diseases.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:132 as residues: Arg-45 to Ser-53.

10

15

20

25

30

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 9

The translation product of this gene shares week sequence homology with rat N-methyl-D-aspartate receptor subunit and other proline-rich proteins which are thought to be important in neurotransmission or protein-protein intereaction.

This gene is expressed primarily in synovial hypoxia and to a lesser extent in ovary, senescent cells and brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, synovial hypoxia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the synovia and brain, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., synovial tissue, ovary and other reproductive tissue, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in synovial hypoxia and nerve tissues, and homology to N-methyl-D-aspartate receptor subunit and other proline-rich proteins indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and intervention of synovial hypoxia and other synovial disorders.

10

15

20

25

30

35

FEATURES OF PROTEIN ENCODED BY GENE NO: 10

This gene is expressed primarily in prostate and to a lesser extent in placenta and ovary.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, male and female infertility, cancer, and other hyperproliferative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of these tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system and neoplasia, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., prostate, placenta, ovary and other reproductive tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:134 as residues: Pro-17 to Met-23, Ala-30 to Trp-38, Ile-49 to Trp-54, Lys-68 to Gly-74, Thr-93 to Gly-99, Met-126 to Glu-132, Gly-173 to Ser-178, Lys-205 to Tyr-214.

The tissue distribution of this gene in the prostate, placenta and ovary indicates that this gene product is useful for treatment/diagnosis of male or female infertility, endocrine disorders, fetal deficiencies, ovarian failure, amenorrhea, ovarian cancer, benign prostate hyperplasia, prostate cancer, and other forms of cancer of the reproductive system.

FEATURES OF PROTEIN ENCODED BY GENE NO: 11

This gene is expressed primarily in the thyroid and to a lesser extent in the pineal gland. This gene maps to chromosome 10, and therefore can be used as a marker in linkage analysis for chromosome 10. Preferred polypeptide fragments comprise the amino acid sequence: HPIEWAINAATLSQFY (SEQ ID NO:256); CWIKYCLTLMQN AQLSMQDNIG (SEQ ID NO:257); KVSYLRPLDFEEARELFLLGQHYVF (SEQ ID NO:258); MERRCKMHKRXIAMLEPLTVDLNPQ (SEQ ID NO:259); and/or SHIV KKINNLNKSALKY YQLFLD (SEQ ID NO:260). Also preferred are polynucleotides encoding these polypeptide fragments.

Therefore, polynucleotides and polypeptides of the invention are useful as

10

15

20

25

30

35

reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune, thyroid and pineal gland disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of these tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and endocrine systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., thyroid and pineal gland, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:135 as residues: Ser-2 to Ser-8, Thr-38 to Arg-44.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treating/detecting immune disorders such as arthritis, asthma, immune deficiency diseases (e.g., AIDS), and leukemia, as well as treating/detecting thymus disorders (e.g., Graves Disease, lymphocytic thyroiditis, hyperthyroidism, and hypothyroidism), and treating/detecting pineal gland disorders (e.g., circadian rhythm disturbances associated with shift work, jet lag, blindness, insomnia and old age).

FEATURES OF PROTEIN ENCODED BY GENE NO: 12

This gene is expressed primarily in lung and tonsils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, pulmonary or immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of these tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the pulmonary and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., pulmonary tissue, and tonsils, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily

fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:136 as residues: Glu-28 to Gly-49.

The tissue distribution of this gene only in lung indicates that it could play a role in the treatment/detection of lung lymphoma or sarcoma formation, pulmonary edema and embolism, bronchitis and cystic fibrosis. Its expression in tonsils indicates a potential role in the treatment/detection of immune disorders such as arthritis, asthma, immune deficiency diseases (e.g., AIDS), and leukemia, in addition to the treatment/detection of tonsillitis.

FEATURES OF PROTEIN ENCODED BY GENE NO: 13

This gene is expressed primarily in lymphoid, myeloid and erythroid cells. Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, hematopoietic and immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of these tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoietic and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, myeloid cells, and bone marrow, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The predominant tissue distribution of this gene in hematopoietic cell types indicates that the gene could be important for the treatment or detection of immune or hematopoietic disorders including arthritis, asthma, immunodeficiency diseases and leukemia. Preferred embodiments of the present invention are polypeptide fragments comprising the amino acid sequence: FTHLSTCLLSLLLVRMSGFLLLARASPSI CALDSSCFVEYCSSYSSSCFLHQHFPSLLDHLCQ (SEQ ID NO:261); or FLLL ARASPSICALDSSCFVQEY (SEQ ID NO:262). Also preferred are polynucleotide fragments encoding these polypeptide fragments.

FEATURES OF PROTEIN ENCODED BY GENE NO: 14

This gene is homologous to the *Drosophila Regena* (Rga) gene. (See Accession No. 1658504.) This *Drosophila* gene is thought to be a homolog of the global negative

10

15

20

25

30

35

10

15

20

35

transcriptional regulator NOT2 (CDC36) from yeast, which modifies gene expression and suppresses position effect variegation. Preferred polypeptide fragments comprise the amino acid sequence: PDGRVTNIPQGMVTDQFGMIGLLTFIRAAETDPGMVHL ALGSDLTTLGLNLNS (SEQ ID NO:263); VHLALGSDLTTLGLNLNSPENLYP (SEQ ID NO:265); EDLLFYLYYMNGGDVLQLLAAVELFNRDWRYHKEERVWI TR (SEQ ID NO:264); and/or HNEDFPALPGS (SEQ ID NO:266).

This gene is expressed primarily in placenta and to a lesser extent in infant brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, neurodegenerative and developmental disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the neurological system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., placenta, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:138 as residues: Leu-9 to Tyr-15, Asp-34 to Gln-46, Pro-51 to Asp-57, Gly-88 to Thr-104, Thr-123 to Ser-128.

The tissue distribution of this gene indicates that it could be used in the detection and/or treatment of neurological disorders such as such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, and panic disorder.

30 FEATURES OF PROTEIN ENCODED BY GENE NO: 15

This gene is expressed primarily in adrenal gland tumor and osteoclastoma. Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, endocrine and bone disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for

10

15

20

25

30

35

differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine system and in bone, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., adrenal gland, and bone, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:139 as residues: Ile-52 to Trp-57.

The tissue distribution of this gene indicates that it may be involved in the treatment and/or detection of adrenal gland tumors, osteosarcomas, endocrine disorders and bone disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 16

The translation product of this gene shares sequence homology with the FK506 binding protein, a protein which plays an important role in immunosupression. (See Accession No. M75099.) Specifically, a 12-kDa FK506-binding protein (FKBP-12) is a cytosolic receptor for the immunosuppressants FK506 and rapamycin. (See, Proc. Natl. Acad. Sci. 88: 6677-6681 (1991).) Thus, based on homology, it is likely that this gene also has immunosuppression activity. Preferred polypeptides comprise the amino acid sequence: GRIIDTSLTRDPLVIELGQKQVIPGLEQSLLDMCVGEKRRAIIPSH LAYGKRGFPPSVPADAVVQYDVELIALIR (SEQ ID NO:267); and/or IHYTGSLV DGR IIDTS (SEQ ID NO:268). Also preferred are the polynucleotide fragments encoding these polypeptides.

This gene is expressed primarily in melanocytes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, cancer and other hyperproliferative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and cancer, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., melanocytes, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to

15

20

25

30

35

the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:140 as residues: Ala-118 to Phe-124, Arg-178 to Lys-201.

The tissue distribution and homology to the FK506 binding proteins which are believed to a role in immunosupression mediated by the immunosupressant drugs rapamycin and cyclosporin, indicates that this gene could serve as a novel target for the identification of novel immunosupressant drugs.

10 FEATURES OF PROTEIN ENCODED BY GENE NO: 17

The translation product of this gene shares sequence homology with the rat calcium-activated potassium channel rSK3, which is thought to be important in regulating vascular tone. (See Accession No. gil2564072, gil1575663, and gil1575661.) Although homologous to these proteins, this gene contains an 18 amino acid insert, not previously identified in the homologs. Preferred polypeptide fragments comprise the amino acid sequence: CESPESPAQPSGSSLPAWYH (SEQ ID NO:269). Also preferred are the polynucleotide fragments encoding these polypeptides.

This gene is expressed primarily in B-cells, frontal cortex and endothelial cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, cardiovascular (hyper/hypotension, asthma, pulmonary edema, pneumonia, heart disease, restenosis, atherosclerosis, stoke, angina and thrombosis) and neurological disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cardiovascular and nervous systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, brain and other tissue of the nervous system, and endothelium, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:141 as residues: Glu-72 to Gly-82, His-90 to Val-95, Gln-168 to Lys-174, Val-202 to Ser-212.

10

15

20

25

30

The tissue distribution and homology to calcium-activated potassium channels indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of vascular disorders (hyper/hypotension, athesma, pulmonary edema, pneumonia, heart disease, restenosis, atherosclerosis, stoke, angina and thrombosis).

FEATURES OF PROTEIN ENCODED BY GENE NO: 18

This gene is expressed primarily in smooth muscle and to a lesser extent in brain (amygdala, corpus colosum, hippocampus).

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, cardiovascular (hypertension, heart disease, athesma, pulmonary edema, restenosis, atherosclerosis, stoke, angina, thrombosis, and wound healing), and neurological disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cardiovascular and neurological systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., smooth muscle, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:142 as residues: Lys-43 to Arg-49, Tyr-58 to Glu-65.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of cadiovascular disorders (hypertension, heart disease, athesma, pulmonary edema, restenosis, atherosclerosis, stoke, angina, thrombosis, and wound healing). Expression in brain indicates a role in the treatment and diagnosis of behavioral or neurological disorders, such as depression, schizophrenia, Alzheimer's disease, mania, dementia, paranoia, and addictive behavior.

35 FEATURES OF PROTEIN ENCODED BY GENE NO: 19

This gene is expressed primarily in T-cells (Jurkats, resting, activated, and

10

15

20

25

30

35

anergic T-cells), endothelial cells, pineal gland, and to a lesser extent in a variety of other tissues and cell types. Preferred polypeptide fragments comprise the amino acid sequence: EEAGAGRRCSHGGARPAGLGNEGLGLGGDPDHTDTGSRSKQRINN WKESKHKVIMASASARGNQDKDAHFPPPSKQSLLFCPKSKLHIHRAEISK (SEQ ID NO:270); and/or SKQRINNWKESKHKVIMASASAR (SEQ ID NO:271). Also preferred are the polynucleotide fragments encoding these polypepides.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, inflammation, immune and cardiovascular disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of these tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, neurological and vascular systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., T-cells and other blood cells, endothelial cells, and pineal gland, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:143 as residues: Phe-71 to Arg-76, Pro-82 to His-87, Glu-103 to Ala-111.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of immune disorders including: leukemias, lymphomas, auto-immune, immuno-supressive (e.g. transplantation) and immunodeficiencies (e.g. AIDS) and hematopoietic disorders. In addition, expression in the pineal gland might suggest a role in the diagnosis of specific brain tumors and treatment of neurological disorders. Endothelial cell expression might suggest a role in cadiovascular or respiratory/pulmonary disorders or infections (athesma, pulmonary edema, pneumonia).

FEATURES OF PROTEIN ENCODED BY GENE NO: 20

This gene is expressed primarily in brain and embryo and to a lesser extent in leukocytes. This gene maps to chromosome 15, and therefore can be used as a marker in linkage analysis to chromosome 15.

Therefore, polynucleotides and polypeptides of the invention are useful as

10

15

20

25

30

35

reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, developmental and neurological disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:144 as residues: Met-1 to Gly-8.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of immune disorders including: leukemias, lymphomas, auto-immune, immuno-supressive (e.g. transplantation) and immunodeficiencies (e.g. AIDS) and hematopoietic disorders. The expression in the brain -- and in particular the fetal brain -- would suggest a possible role in the treatment and diagnosis of developmental and neurodegenerative diseases of the brain and nervous system (depression, schizophrenia, Alzheimer's disease, mania, dementia, paranoia, and addictive behavior).

FEATURES OF PROTEIN ENCODED BY GENE NO: 21

This gene is expressed primarily in brain, kidney, lung, liver, spleen, and a variety of leukocytes (especially T-cells) and to a lesser extent in a variety of other tissues and cell types.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, leukemias, lymphomas, autoimmune, immunosuppressive, and immunodeficiencies, hematopoietic disorders, as well as renal disorders, and neoplasms. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the renal, pulmonary, immune, and central nervous systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.,

10

15

20

25

30

35

brain and other tissue of the nervous system, kidney, pulmonary tissue, liver, spleen, and blood cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of renal conditions, such as acture renal failure, kidney fibrosis, and kidney tubule regeneration. The expression in leukocytes and other immune tissues indicates a role in immune disorders including: leukemias, lymphomas, auto-immune, immuno-supressive (e.g. transplantation) and immunodeficiencies (e.g. AIDS) and hematopoietic disorders. The expression in the brain -- and in particular the fetal brain -- indicates a possible role in the treatment and diagnosis of developmental and neurodegenerative diseases of the brain and nervous system (depression, schizophrenia, Alzheimer's disease, mania, dementia, paranoia, and addictive behavior).

FEATURES OF PROTEIN ENCODED BY GENE NO: 22

This gene is expressed primarily in skin (fetal epithelium, keratinocytes and skin). This gene also maps to chromosome 19, and therefore can be used in linkage analysis as a marker for chromosome 19.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, skin cancers (e.g., melanomas), eczema, psoriasis or other disorders of the skin. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of these tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skin, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., skin and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:146 as residues: Pro-28 to Glu-35, Ser-39 to Phe-44, Ala-94 to Gln-99.

15

20

25

30

35

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of skin cancers (e.g., melanomas), eczema, psoriasis or other disorders of the skin.

FEATURES OF PROTEIN ENCODED BY GENE NO: 23

This gene maps to chromosome 11. Another group recently isolated this same gene, associating the sequence to the region thought to harbor the gene involved in Multiple Endocrine Neoplasia Type 1, or MEN 1. (See Accession No. 2529721 and Genome Res. 7(7), 725-735 (1997), incorporated herein by reference in its entirety.) Preferred polypeptide fragments comprise the amino acid sequence: LFHWACLNERA AOLPRNTAXAGYQCPSCNGPS (SEQ ID NO:272).

This gene is expressed primarily in epididymus, pineal gland, T-cells, as well as fetal epithelium, lung and kidney.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune, metabolic mediated disorders, and MEN. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, renal, neurological and pulmonary systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., epididymus and other reproductive tissue, pineal gland, T-cells and other blood cells, epithelium, lung, and kidney, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of developmental deficiencies or abnormalities as well as a host of different disorders which arise as a result of conditions in the indicated tissues or cell types. An area of particular interest is in the treatment and diagnosis of immune disorders including: leukemias, lymphomas, auto-immune, immuno-supressive (e.g. transplantation) and immunodeficiencies (e.g. AIDS) and hematopoietic disorders. The expression in the brain, and in particular the fetal brain, would suggest a possible role in the treatment and diagnosis of

10

15

20

25

30

35

developmental and neurodegenerative diseases of the brain and nervous system (depression, schizophrenia, Alzheimer's disease, mania, dementia, paranoia, and addictive behavior). Respiratory/pulmonary disorders, such as athesma, pulmonary edema are also potential therapeutic areas, as well as renal conditions such as acute renal failure, kidney fibrosis and kidney tubule regeneration. Moreover, this gene can be used in the treatment and/or detection of MEN I.

FEATURES OF PROTEIN ENCODED BY GENE NO: 24

This gene is expressed primarily in fetal spleen.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, leukemia, lymphoma, AIDS, hematopoeitic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hematopoietic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., spleen and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of immune disorders including: leukemias, lymphomas, auto-immune, immuno-supressive (e.g. transplantation) and immunodeficiencies (e.g. AIDS) and hematopoietic disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 25

A closely related homolog of this gene was recently cloned by another group, calling the gene CDO, an oncogene-, serum-, and anchorage-regulated member of the Ig/fibronectin type III repeat family. (See Accession No. 2406628, and J. Cell Biol. 138(1): 203-213 (1997), herein incorporated by reference in its entirety.) Preferred polypeptide fragments comprise the amino acid sequence: FYIYYRPTDSDNDSDYKK DMVEGDKYWHSISHLQPETSYDIKMQCFNEGGESEFSNVMICETKARKSSGQP GRLPPPTLAPPQPPLPETIERPVGTGAMVARSSDLPYLIVGVVLGSIVLIIVTFIPF CLWRAWSKQKHTTDLGFPRSALPPSCPYTMVPLGGLPGHQAVDSPTSVASVD

10

15

20

25

30

35

GPVLM (SEQ ID NO:273); or YTYYRPTDSDNDSDYKKDMVEGDKYWHSISHLQ PETSYDIKMQCFNEGGESEFSNVMICETKARKS (SEQ ID NO:274).

This gene is expressed primarily in fetal lung and kidney, human embryo and osteoclastoma stromal cells and to a lesser extent in a variety of other tissues and cell types.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, developmental disorders and cancers, as well as pulmonary and renal disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the respiratory/pulmonary, skeletal and renal systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., lung, kidney, embryonic tissue, and bone cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:149 as residues: Thr-5 to Pro-18, Ala-76 to Thr-84.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of: osteoperosis, fracture, osteosarcoma, ossification, and osteonecrosis, as well as respiratory/pulmonary disorders, such as athesma, pulmonary edema, and renal conditions such as acute renal failure, kidney fibrosis and kidney tubule regeneration.

FEATURES OF PROTEIN ENCODED BY GENE NO: 26

This gene is homologous to the HIV envelope glycoprotein. (See Accession No. 2641463.) Preferred polypeptide fragments comprise the amino acid sequence: NVRALLHRMPEPPKINTAKFNNNKRKNLSL (SEQ ID NO:275).

This gene is expressed primarily in pineal gland and skin, and to a lesser extent in lung.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are

not limited to, neurological and behavior disorders; respiratory/pulmonary disorders, such as athesma, pulmonary edema; skin conditions such as eczema, psoriasis, acne and skin cancer, as well as AIDS. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous and respiratory systems, as well as skin and AIDS, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, pineal gland, epidermis, and pulmonary tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:150 as residues: Gln-15 to Gln-20.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of conditions which affect the above tissues, such as: skin cancer, eczema, psoriasis, acne, athesma, pulmonary edema, neuro-degenerative or developmental disorders such as Alzheimer's, depression, schizophrenia, dementia, and AIDS.

20

25

30

35

DEICHOON AND DOINTSOAT IA.

15

5

10

FEATURES OF PROTEIN ENCODED BY GENE NO: 27

Preferred polypeptide encoded by this gene comprise the following amino acid sequence: NTNQREALQYAKNFQPFALNHQKDIQVLMGSLVYLRQGIENSPYVHL LDANQWADICDIFTRDACALLGLSVESPLSVSFSAGCVALPALINIKAVIEQRQC TGVWNQKDELPIEVDLGKKCWYHSIFACPILRQQTTDNNPPMKLVCGHIISRD ALNKMFNGSKLKCPYCPMEQSPGDAKQIFF (SEQ ID NO:276). Polynucleotides encoding such polypeptides are also provided as are complementary polynucleotides thereto.

This gene is expressed primarily in liver (adult and fetal) and spleen tissue, and to a lesser extent in placenta, T helper cells, kidney tumor, ovarian tumor, melanocytes and fetal heart.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune and developmental diseases and disorders and liver diseases such as liver cancer. Similarly, polypeptides and antibodies directed to these

10

20

25

30

35

polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, circulatory and hematopoietic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., liver, spleen, placenta, blood cells, kidney, ovary and other reproductive tissue, melanocytes, and heart, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that the protein products of this gene are useful for study, diagnosis and treatment of growth, hematopoietic and immune system disorders particularly related to the liver.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 28

The translation product of this gene shares sequence homology with prostaglandin transporter which is thought to be important in metabolic and endocrine disorders. See, for example, Gastroenterology Oct:109(4):1274-1282 (1995). Preferred polypeptides encoded by this gene comprise the following amino acid sequence: SYLSACFAGCNSTNLTGCACLTTVPAENATVVPGKCPSPGCQEAFLTFLCVMCI CSLIGAMARHP (SEQ ID NO:277); and/or PSVIILIRTVSPELKSYALGVLFLLLRL LGFIPPPLIFGAGIDSTCLFWSTFCGEQGACVLYDNVVYRYLYVSIAIALKSFAFI (SEQ ID NO:278).

This gene is expressed primarily in hematopoietic and brain tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, metabolic, immune and endocrine diseases and disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the metabolic, immune and endocrine systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., endocrine tissue, hematopoietic tissue, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to

10

15

20

25

30

35

the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to prostaglandin (and anion) transporter indicates that polynucleotides and polypeptides corresponding to this gene are useful for study, diagnosis and treatment of endocrine, metabolic, immune and kidney disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 29

This gene is expressed primarily in early stage human lung.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, growth and respiratory disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the developmental and respiratory systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., pulmonary tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:153 as residues: Val-50 to Trp-55.

The tissue distribution indicates that the protein products of this gene are useful for study, diagnosis and treatment of respiratory and growth diseases and disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 30

The translation product of this gene shares sequence homology with human DNA helicase which is thought to be important in accurate and complete DNA replication in creation of new cells. Preferred polypeptides encoded by this gene comprise the following amino acid sequence: QSLFTRFVRVGVPTVDLDAQGRARA SLCXXYNWRYKNLGNLPHVQLLPEFSTANAGLLYDFQLINVEDFQGVGESEPN PYFYQNLGEAEYVVALFMYMCLLGYPADKISILTTYNGQKHLIRDIINRRCGNN PLIGRPNKVTTVDRFQGQQNDYILLSLVRTRAVGHLRDVRRLVVAMSRAR (SEQ ID NO:279); and/or LVKEAKIIAMTCTHAALKRHDLVKLGFKYDNILMEE AAQILEIETFIPLLLQNPQDGFSRLKRWIMIGDHHQLPPVI (SEQ ID NO:280).

This gene is expressed primarily in testes tumor and to a lesser extent in adrenal

gland tumor and placenta.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, cancers and endocrine/growth disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine, developmental, and reproductive systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., testes and other reproductive tissue, adrenal gland, and placenta, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to DNA helicase indicates that the protein products of this gene are useful for study, treatment, and diagnosis of many cancer types, including testicular cancer; as well as disorders involving endocrine function and normal growth and development.

20

5

10

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 31

The translation product of this gene shares sequence homology with BID-apoptotic death gene (mouse), Genbank accession no. PID g1669514, which is thought to be important in programmed cell death.

25

This gene is expressed primarily in jurkat membrane bound polysomes and activated neutrophils and to a lesser extent in endothelial cells and human cerebellum.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, cancers and other proliferative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, endothelium, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma,

35

30

10

15

20

25

30

35

urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:155 as residues: Glu-4 to Leu-11, Cys-28 to Arg-35, Gln-50 to His-66, Glu-73 to Gln-79, Gly-94 to Ser-100, Arg-114 to Asp-126, Pro-139 to Lys-146.

The tissue distribution and homology to BID-apoptotic death gene indicates that the protein products of this gene are useful for study of cell death, and treatment and diagnosis of proliferative disorders and cancers. Apoptosis - programmed cell death - is a physiological mechanism involved in the deletion of peripheral T lymphocytes of the immune system, and its dysregulation can lead to a number of different pathogenic processes. Diseases associated with increased cell survival, or the inhibition of apoptosis, include cancers (such as follicular lymphomas, carcinomas with p53 mutations, and hormone-dependent tumors, such as breast cancer, prostrate cancer, Kaposiís sarcoma and ovarian cancer); autoimmune disorders (such as systemic lupus erythematosus and immune-related glomerulonephritis rheumatoid arthritis) and viral infections (such as herpes viruses, pox viruses and adenoviruses), inflammation; graft vs. host disease, acute graft rejection, and chronic graft rejection. Diseases associated with increased apoptosis include AIDS; neurodegenerative disorders (such as Alzheimer's disease, Parkinson's disease, Amyotrophic lateral sclerosis, Retinitis pigmentosa, Cerebellar degeneration); myelodysplastic syndromes (such as aplastic anemia), ischemic injury (such as that caused by myocardial infarction, stroke and reperfusion injury), toxin-induced liver disease (such as that caused by alcohol), septic shock, cachexia and anorexia. Thus, the invention provides a method of enhancing apoptosis in an individual by treating the individual with a polypeptide encoded by this gene.

FEATURES OF PROTEIN ENCODED BY GENE NO: 32

The translation product of this gene shares sequence homology with human fructose transporter which is thought to be important in normal metabolic function and activity.

This gene is expressed primarily in T-cell lymphoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are

10

15

20

25

30

35

not limited to, leukemia and other cancers, and metabolic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoietic, lymph and metabolic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., T-cells and other blood cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:156 as residues: Pro-22 to Gly-48, Ser-54 to Pro-61.

The tissue distribution indicates that the protein products of this gene are useful for study of mechanisms leading to cancer, treatment and diagnosis of cancerous and pre-cancerous conditions; as well as the study and treatment of various metabolic diseases and disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 33

This gene is expressed primarily in human meningima.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, inflammation and other disorders of the CNS. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the CNS and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., meningima and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:157 as residues: Asn-23 to Pro-31.

The tissue distribution indicates that the protein products of this gene are useful for study, diagnosis and treatment of disorders of the CNS and inflammatory responses.

10

15

20

30

35

FEATURES OF PROTEIN ENCODED BY GENE NO: 34

This gene is expressed primarily in activated monocytes and wound healing tissues and to a lesser extent in fetal epithelium.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune and inflammatory disorders and wound healing and tissue repair dysfunctions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, epithelial and gastrointestinal systems, and healing wounds, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., monocytes and other blood cells, and epithelium, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:158 as residues: Ala-28 to Ala-33, Gly-35 to Glu-45.

The tissue distribution indicates that the protein products of this gene are useful for diagnosis, study and treatment of immune and inflammatory disorders and wound healing dysfunctions.

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 35

This gene is expressed primarily in human osteosarcoma and prostate cancer.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, skeletal and neoplastic conditions such as bone and prostate cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and skeletal systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., bone, and prostate, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial

10

15

20

25

30

35

fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:159 as residues: Ser-14 to Gly-22, Leu-37 to Gln-43.

The tissue distribution indicates that the protein products of this gene are useful for diagnosis and treatment of skeletal disorders and cancer.

FEATURES OF PROTEIN ENCODED BY GENE NO: 36

This gene encodes a protein which is highly homologous to a protein called congenital heart disease protein 5, presumably implicated in congenital heart disease (see Genbank PID g2810996).

This gene is expressed primarily in Hodgkin's lymphoma, erythroleukemia cells, and TNF activated synovial fibroblasts, to a lesser extent in ovarian cancer, cerebellum, spleen, fetal liver and placenta and finally to a lesser extent in various other mesenchymal tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, cancer, immune, hematopoietic and cardiovascular disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, hematopoietic and cardiovascular systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., heart and other cardiovascular tissue, lymphoid tissue, blood cells, bone marrow, ovary and other reproductive tissue, brain and other tissue of the nervous system, spleen, liver, and mesenchymal tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:160 as residues: Lys-41 to Met-49, Gln-54 to Glu-59, Glu-76 to Thr-88.

The homology of this gene and translation product to congenital heart disease protein 5 indicates a role for this protein in the diagnosis, prognosis and/or treatment of

heart disease or other cardiovascular related disorders. In addition, predominant expression in cells associated with the immune and hematopoetic system indicates a role for this protein in the treatment, diagnosis and/or prognosis of immune and autoimmune diseases, such as lupus, transplant rejection, allergic reactions, arthritis, asthma, immunodeficiency diseases, leukemia, AIDS, thymus disorders such as Graves Disease, lymphocytic thyroiditis, hyperthyroidism and hypothyroidism, graft versus host reaction, graft versus host disease, transplant rejection, myelogenous leukemia, bone marrow fibrosis, and myeloproliferative disease. The protein could also be used to enhance or protect proliferation, differentiation and functional activation of hematopoietic progenitor cells such as bone marrow cells, which could be useful for cancer patients undergoing chemotherapy or patients undergoing bone marrow transplantation. The protein may also be useful to increase the proliferation of peripheral blood leukocytes, which could be useful in the combat of a range of hematopoietic disorders including immunodeficiency diseases, leukemia, and septicemia.

15

20

25

10

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 37

This gene is expressed primarily in ovarian cancer.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, urogenital neoplasias. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., ovary and other reproductive tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:161 as residues: Asn-22 to Asn-27.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for study, diagnosis and treatment of ovarian and other tumors.

30

10

15

20

25

30

35

FEATURES OF PROTEIN ENCODED BY GENE NO: 38

The translation product of this gene shares sequence homology with zinc finger proteins.

This gene is expressed primarily in various fetal, cancer, and endothelial lines.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune and growth disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cardiovascular system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., fetal tissue, and endothelial cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that the protein products of this gene are useful for study, diagnosis and treatment of immune and developmental conditions and cancer.

FEATURES OF PROTEIN ENCODED BY GENE NO: 39

This gene is expressed primarily in fetal, infant, and adult brain and to a lesser extent in other brain and endocrine organs and blastomas.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, brain tumors and neurodegenerative conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous and endocrine systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., brain and other tissue of the nervous system, endocrine tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the

disorder.

10

15

20

25

30

The tissue distribution indicates that the protein products of this gene are useful for the study, diagnosis and treatment of brain cancer and other neurological disorders.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 40

The translation product of this gene shares sequence homology with vesicular glycoproteins and lectins. Preferred polypeptides encoded by this gene comprise the following amino acid sequence: DTYPNEEKQQERVFPXXSAMVNNGSLSYDHER DGRPTELGGCXAIVRNLHYDTFLVIRYVKRHLTIMMDIDGKHEWRDCIEVPGV RLPRGYYFGTSSITGDLSDNHDVISLKLFELTVERTPEEE (SEQ ID NO:281); and/or LKREHSLSKPYQGVGTGSSSLWNLMGNAMVMTQYIRLTPDMQSKQGA LWNRVPCFLRDWELQVHFKIHGQGKKNLHGDGLAIWYT (SEQ ID NO:282).

This gene is expressed primarily in infant brain and to a lesser extent in various normal and transformed neural, endocrine, and immune organs.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, neurological and neurodevelopmental conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous and hormonal systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., brain and other tissue of the nervous system, endocrine tissue, and tissue and cells of the immune system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:164 as residues: Pro-64 to Gly-71, Gly-94 to Leu-100, Thr-110 to Pro-116, Thr-135 to Arg-145, Glu-164 to Glu-171, Asp-204 to Asp-211, Arg-253 to His-261, Asn-312 to Tyr-323.

The tissue distribution indicates that the protein products of this gene are useful for the study, diagnosis and treatment of mental retardation and other neurological disorders and neoplasias.

10

15

20

FEATURES OF PROTEIN ENCODED BY GENE NO: 41

This gene displays homology to the glycosyltransferase family, which catalyze the addition of sialic acids to carbohydrate groups which are present on glycoproteins.

This gene is expressed primarily in smooth muscle and to a lesser extent in pineal gland, fetal liver, and infant brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, gastrointestinal injury, inflammatory and neurodegenerative conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and nervous systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., smooth muscle, pineal gland, liver, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:165 as residues: Ser-12 to Trp-21, Arg-24 to Pro-32, Asp-73 to Lys-82, Lys-90 to Ala-97.

The tissue distribution indicates that the protein products of this gene are useful for the study, diagnosis and treatment of neurodegenerative and growth disorders and gastrointestinal repair.

25

30

35

FEATURES OF PROTEIN ENCODED BY GENE NO: 42

The translation product of this gene shares sequence similarity with metallothionein polypeptides. See, for example, Proc. Natl. Acad. Sci. U S A 1992 Jul 15:89(14):6333-6337. Metallothioneins are believed to inhibit neuronal survival among other biological functions. Based on the sequence similarity (especially the conserved cysteine motifs characteristic of the metallothionein family) the translation product of this gene is expected to share certain biological activities with other members of the metallothionein polypeptide family. Preferred polypeptides encoded by this gene comprise the following amino acid sequence: PGTLQCSALHHDPGCANCSRFCRD CSPPACQC (SEQ ID NO:283).

This gene is expressed exclusively in placenta and fetal liver.

10

15

25

30

35

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, hematopoietic and immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., placenta, liver, brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to metallothionien indicates that the protein products of this gene are useful for diagnosis and treatment of immune and hematopoietic system disorders and neurological diseases, especially in fetal development.

20 FEATURES OF PROTEIN ENCODED BY GENE NO: 43

Preferred polypeptides encoded by this gene comprise the following amino acid sequence: FLYDVLMXHEAVMRTHQIQLPDPEFPS (SEQ ID NO:284).

This gene is expressed primarily in T-cells and synovial tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune system disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., synovial tissue, and T-cells and other blood cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

10

15

20

25

30

35

The tissue distribution indicates that the protein products of this gene are useful for treatment and diagnosis of disorders of the immune system.

FEATURES OF PROTEIN ENCODED BY GENE NO: 44

The translation product of this gene shares sequence similarity with several methyltransferases (e.g., see Genbank gill065505).

This gene is expressed primarily in ovary, thymus, infant adrenal gland, tissues of the nervous system and the hematopoietic tissue, and to a lesser extent in adipose tissue and many other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, disorders of the reproductive system, the endocrine system, the hematopoietic system and the CNS. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, endocrine, CNS and reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., ovary and other reproductive tissue, thymus, adrenal gland, brain and other tissue of the nervous system, hematopoietic tissue, and adipose tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:168 as residues: Ser-3 to Gly-12, Asp-19 to Arg-31, Tyr-70 to Tyr-77, Asn-130 to Lys-140, Pro-165 to Gln-170, Pro-192 to Lys-199, Leu-216 to Glu-227, Glu-254 to Phe-281.

The tissue distribution and homology to methyltransferase indicates that the protein products of this gene are useful for diagnosis and treatment of disorders of the CNS, the hematopoietic system and reproductive organs and tissues. For example, the abundant expression in the ovary may indicate that the gene product can be used as a hormone with either systemic or reproductive functions; as growth factors for germ cell maintenance and in vitro culture; as a fertility control agent; remedy for sexual dysfunction or sex development disorders; diagnostics/treatment for ovarian tumors, such as serous adenocarcinoma, dysgerminoma, embryonal carcinoma,

10

15

20

25

30

35

choriocarcinoma, teratoma, etc; The expression in thymus may indicate its utilities in T-cell development and thus its applications in immune related medical conditions, such as infection, allergy, immune deficiency, tissue/organ transplantation, etc.

FEATURES OF PROTEIN ENCODED BY GENE NO: 45

The translation product of this gene shares sequence homology with cytochrome C oxidase which is thought to be important in metabolic function of cells. This gene has now recently been published as estrogen response gene. See Genbank accession no. AB007618 and Mol. Cell. Biol. 18 (1), 442-449 (1998). See also J Immunol. Mar 1:154(5): 2384-2392 (1995), where the mouse homologue was published and implicated in siliocis.

This gene is expressed primarily in adipose tissue, kidney and fetal brain and to a lesser extent in several other tissues and organs.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, metabolic diseases involving especially adipose tissue, brain and kidney. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the CNS and vascular system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., adipose tissue, kidney, brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:169 as residues: Thr-5 to Ser-14.

The tissue distribution and homology to cytochrome C oxidase, estrogen response gene product and siliocis related gene product indicates that the protein products of this gene are useful for diagnosis and treatment of metabolic disorders in the CNS, adipose tissue and kidney, particularly siliocis.

FEATURES OF PROTEIN ENCODED BY GENE NO: 46

The translation product of this gene shares sequence homology with reticulocalbin. See, for example, J. Biochem. 117 (5), 1113-1119 (1995). Based on the

10

15

20

25

30

35

sequence similarity, the translation product of this gene is expected to share certain biological activities with reticulocalbin, e.g., Ca++ binding activities. This gene product is sometimes hereinafter referred to as "Reticulocalbin-2".

This gene is expressed primarily in breast, endothelial cells, synovial, heart and smooth muscle cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, diseases of the breast, vascular and skeletal/cardiac muscular system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the breast, vascular and skeleto-muscular system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., mammary tissue, endothelial cells, synovial tissue, heart and other cardiovascular tissue, and smooth muscle, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:170 as residues: Gly-16 to Arg-32, Ala-42 to Asn-50, Glu-66 to Gln-76, Arg-85 to Gly-94, Thr-108 to Asp-115, Trp-121 to Gly-130, Leu-137 to His-144, Glu-155 to Lys-161, Asp-175 to Ser-180, Glu-209 to Gly-217, Glu-232 to Glu-237, Thr-243 to Asp-261, Glu-287 to Arg-295.

The tissue distribution indicates that the protein products of this gene are useful for diagnosis and treatment of diseases of the vascular and skeletal/cardiac muscular system. The homology of the gene with reticulocalbin indicates its biological function in regulating calcium store, a particularly important function in muscular cell types. The gene expression in the heart may indicate its utilities in diagnosis and remedy in heart failure, ischemic heart diseases, cardiomyopathy, hypertension, arrhythmia, etc. The abundant expression in the breast may indicate its applications in breast neoplasia and breast cancers, such as fibroadenoma, papillary carcinoma, ductal carcinoma, Pagetís disease, medullary carcinoma, mucinous carcinoma, tubular carcinoma, secretory carcinoma and apocrine carcinoma; juvenile hypertrophy and gynecomastia, mastitis and abscess, duct ectasia, fat necrosis and fibrocystic diseases, etc.

10

15

20

FEATURES OF PROTEIN ENCODED BY GENE NO: 47

The translation product of this gene shares weak sequence homology with H+transporting ATP synthase which is thought to be important in cell metabolism or signal transduction.

This gene is expressed only in testis.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of some types of diseases and conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain and hematopoietic tissues, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., testes and other reproductive tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Since only one out of about a million expressed sequence tag is found in testes indicates that its expression is low and selectively in testes. Since some of the genes only expressed in testes are usually expressed in brain or in certain induced hematopoietic cells/tissues, it is speculated that this gene to be expressed in brain or hematopoietic cells/tissues and is useful for diagnosis and treatment of disorders these systems.

25

30

35

FEATURES OF PROTEIN ENCODED BY GENE NO: 48

The translation product of this gene shares sequence homology with human polymeric immunoglobulin receptor (accession No.X73079) which is thought to be important in antibody recognition and immune defenses. In one embodiment, polypeptides of the invention comprise the sequence GWYWCG (SEQ ID NO:285). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in placenta and to a lesser extent in corpus callosum and fetal liver and spleen.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are

10

15

20

25

30

35

not limited to, disorders of the immune system, e.g. autoimmune diseases and immunodeficiency. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., placenta, liver, and spleen, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:172 as residues: Tyr-37 to Cys-49, Gly-51 to Tyr-56, Lys-88 to Trp-93, Leu-130 to Glu-136.

The tissue distribution and homology to human polymeric immunoglobulin receptor indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of immune disorders, e.g. autoimmune diseases and immunodeficiencies.

FEATURES OF PROTEIN ENCODED BY GENE NO: 49

This gene is expressed in thymus.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune disorder. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., thymus and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of immune disorders, e.g. autoimmunity and immunodeficiency.

10

15

20

25

35

FEATURES OF PROTEIN ENCODED BY GENE NO: 50

Preferred polypeptide encoded by this gene comprise the following amino acid sequence: MKVGARIRVKMSVNKAHPVVSTHWRWPAEWPQMFLHLAQEPRTE VKSRPLGLAGFIRQDSKTRKPLEQETIMSAADTALWPYGHGNREHQENELQKY LQYKDMHLLDSGQSLGHTHTLQGSHNLTALNI (SEQ ID NO:286). Polynucleotides encoding this polypeptide are also provided as are complementary polynucleotides thereto.

This gene is expressed primarily in adrenal gland, pituitary, T helper cells, and breast cells and to a lesser extent in a wide variety of tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the some diseases and conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and endocrine systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., adrenal gland, pituitary, T-cells and other blood cells, and mammary tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:174 as residues: Gln-39 to Ser-47, Arg-57 to Glu-67, Tyr-82 to Gln-95.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of a wide range of disorders, such as immune and endocrine disorders.

30 FEATURES OF PROTEIN ENCODED BY GENE NO: 51

The translation product of this gene shares sequence homology with human Sop2p-like protein which is important in cytoskeleton structure. In one embodiment, polypeptides of the invention comprise the sequence SLHKNSVSQISVLSGGKAKCS QFCTTGMDGGMSIWDVKSLESALKDLKI (SEQ ID NO:287). Polynucleotides encoding this polypeptide are also encompassed by the invention. This gene maps to chromosome 7. Therefore, polynucleotides of the invention can be used in linkage

10

15

20

analysis as a marker for chromosome 7.

This gene is expressed primarily in immune and hematopoietic tissues/cells and to a lesser extent in other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immunological and hematopoietic disorders and inflammation. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hematopoietic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., immune and hematopoietic tissue/cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:175 as residues: Lys-49 to Gln-54, Ala-61 to Arg-66, Lys-82 to Lys-87, Glu-126 to Val-133, His-136 to Ile-141, Glu-175 to Ser-187, Asp-286 to Leu-296, Ala-298 to Ser-310.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of immunological, hematopoietic, and inflammatory disorders, e.g, immunodeficiency, autoimmunity, inflammation.

25

30

35

FEATURES OF PROTEIN ENCODED BY GENE NO: 52

The translation product of this gene shares sequence homology with *Caenorhabditis elegans* R53.5 gene encoding a putative secreted protein without known function.

This gene is expressed primarily in endothelial cells, brain and several highly vascularized, and tumor tissues and to a lesser extent in other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, aberrant angiogensis and tumorigenesis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes

10

15

20

25

30

35

for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the vascular and brain system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., endothelial cells, brain and other tissue of the nervous system, and vascular tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:176 as residues: Thr-43 to Asn-60, Thr-106 to Phe-115, Asp-122 to Arg-133, Arg-186 to Asp-192, Leu-211 to Lys-216.

The tissue distribution and homology to a *C. elegans* secreted protein indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis or treatment of disorders in vascular or brain system, e.g. aberrant angiogenesis, ischemia, neurodegeneration, etc.

FEATURES OF PROTEIN ENCODED BY GENE NO: 53

In one embodiment, polypeptides of the invention comprise the sequence EASKSSHAGLDLFSVAACHRF (SEQ ID NO:288). Polynucleotides encoding this polypeptide are also encompassed by the invention.

This gene is expressed primarily in T-cells and to a lesser extent in brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, lymphocytic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the lymphoid system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., T-cells and other blood cells, brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:177 as residues: Pro-3 to Thr-8, Arg-37 to Asp-46.

10

15

20

25

30

35

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis, treatment, and cure of lymphocytic disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 54

The translation product of this gene shares sequence homology with secreted cartilage matrix protein, a major component of the extracellular matrix of nonarticular cartilage which is thought to be important in cartilage structure. In specific embodiments, polypeptides of the invention comprise the sequence: RCKKCTEGPI DLVFVIDGSKSLGEENFEVVKQF (SEQ ID NO:297); VTGIIDSLTISPKAARVGL LQYSTQVH (SEQ ID NO:290); TEFTLRNFNSAKDMKKAVAHMKYM (SEQ ID NO:291); GKGSMTGLALKHMFERSFTQGEGARPF (SEQ ID NO:292); STRVP RAAIVFTDGRAQDDVSEWASKAKANGITMYAVGVGKAIE (SEQ ID NO:293); EELQEIASEPTNKHLFYAEDFSTMDEISEKLKKGICEALEDS (SEQ ID NO:294); TQRLEEMTQRM (SEQ ID NO:295); PQGCPEQPLH (SEQ ID NO:296); and/or YMGKGSMTGLALKHMFERSFT (SEQ ID NO:289). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in placenta, infant brain, prostate, fetal lung and to a lesser extent in endometrium and fetal tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, abnormal placenta and pregnancy, disorder and injury in brain, prostate, and vasculature. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproduction, neuronal, and vascular systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., placenta, brain and other tissue of the nervous system, prostate, lung and endometrium, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to cartilage matrix protein indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis,

10

15

20

25

30

35

treatment, and cure of abnormalities in placenta and pregnancy, disorder and injury in brain, prostate, and vasculature.

FEATURES OF PROTEIN ENCODED BY GENE NO: 55

The translation product of this gene is the human ortholog of bovine and hamster CII-3, a succinate-ubiquinone oxidoreductase complex II membrane-intrinsic subunit, which is thought to be important in mitochondrial electron transport chain during metabolism. In specific embodiments, the polypeptides of the invention compriseMAALLLRHVGRHCLRAHFSPQLCIRNAVPLGTTAKEEMERFWNKNIG SNRPLSPHITIYS (SEQ ID NO:298); VFPLMYHTWNGIRHLMWDLGKGLKIPQL YQSG (SEQ ID NO:299); MAALLLRHVGRHCLRAH (SEQ ID NO:300); VKSLCL GPALIHTAKFAL (SEQ ID NO:301); VFPLMYHTWNGIRHLMWDLGKGL (SEQ ID NO:302).

This gene is expressed in 8-week old early stage human.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, metabolism disorder. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the [insert system where a related disease state is likely, e.g., immune], expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis, treatment, and cure of metabolism disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 56

This gene is expressed primarily in umbilical vein endothelial cells, human ovarian tumor cells, human meningima cells, and human Jurkat membrane bound polysomes. In specific embodiments, polypeptides of the invention comprise the amino acid sequence: RVWDVRPFAPKERCVKIFQGNV (SEQ ID NO:303); HNFEKNLL

10

15

20

RCSWSPDGSKIAAGSADRFVYV (SEQ ID NO:304); and/or WDTTSRRILYKLPG HAGSINEVAFHPDEPI (SEQ ID NO:305). Polynucleotides encoding these polypeptides are also encompassed by the invention.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, inflammation, immune and cardiovascular disorders and urogenital neoplasias. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of these tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, neurological, urogenital, reproductive system and vascular systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, cells, endothelial cells, ovary and other reproductive tissue, meningima, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:143 as residues: Phe-71 to Arg-76, Pro-82 to His-87, Glu-103 to Ala-111.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of immune disorders including: leukemias, lymphomas, auto-immune, immuno-supressive (e.g. transplantation) and immunodeficiencies (e.g. AIDS) and hematopoietic disorders. In addition, expression in ovarian tumor cells suggests that polynucleotides and polypeptides corresponding to this gene are useful for study, diagnosis, and treatment of ovarian tumors, and other tumors and neoplasias. Further, endothelial cell expression suggests a role in cadiovascular or respiratory/pulmonary disorders or infections (athsma, pulmonary edema, pneumonia).

30

35

25

FEATURES OF PROTEIN ENCODED BY GENE NO: 57

The translation product of this gene shares sequence homology with type I collagen. In specific embodiments, the polypeptides of the invention comprise the sequence: GRIPAPAPSVPAGPDSR (SEQ ID NO:309); VRGRTVLRPGLDAEPE LSPE (SEQ ID NO:306); EQRVLERKLKKERKKEERQ (SEQ ID NO:307); ARRSG

10

15

20

AELAWDYLCRWAQKHKNWRFQKTRQTWLLLHMYDSDKVPDEHFSTLLAYLE GLQGR (SEQ ID NO:255); and/or RLREAGLVAQHPP (SEQ ID NO:308).

Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in epididymus, prostate cell line (LNCAP), and pituitary gland; and to a lesser extent in many other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, abnormalities of the epididymus, prostate (especially prostate cancer), and pituitary gland. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the male reproductive system and neuroendocrine system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., epididymus and other reproductive tissue, prostate, and pituitary gland, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to type I collagen, indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of abnormalities of the epididymus, prostate (especially prostate cancer), and pituitary gland.

25

30

35

FEATURES OF PROTEIN ENCODED BY GENE NO: 58

This gene is expressed primarily in the frontal cortex of the brain from a schizophrenic individual.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, schizophrenia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., brain and other tissue of

10

15

20

25

30

35

the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of schizophrenia.

FEATURES OF PROTEIN ENCODED BY GENE NO: 59

The polypeptide encoded by Gene 59 is homologous to human surface 4 integral membrane protein. In specific embodiments, the polypeptides of the invention comprise the sequence: TGCVLVLSRNFVQYACFGLFGIIALQTIAYSILWDLKF LMRN (SEQ ID NO:310); SRSEGKSMFAGVPTMRESSPKQYMQLGGRVLLV LMFMTLLHFDASFFSIVQNIVG (SEQ IDNO:311); GTAEDFADQFLRVTKQYLP HVARLCLISTFLEDGIRMFQWSEQRDYIDTTWNCGYLLAS (SEQ ID NO:312); LMRNESRS (SEQ ID NO:314); ASFLLSRTSWGTA (SEQ ID NO:315); and/or ASFLLSRTSWGTALMIL (SEQ ID NO:313). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in Hodgkin's lymphoma and lung; and to a lesser extent in many other human tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, Hodgkin's lymphoma, tumors or other abnormalities of the lung. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and respiratory systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., lymphoid tissue, and pulmonary tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:183 as residues: Met-20 to Trp-27.

The tissue distribution indicates that polynucleotides and polypeptides

10

15

20

25

30

35

corresponding to this gene are useful for diagnosis and treatment of Hodgkin's lymphoma, tumors or other abnormalities of the lung.

FEATURES OF PROTEIN ENCODED BY GENE NO: 60

This gene is expressed primarily in bone cancer and stomach cancer, and to a lesser extent in many other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, bone cancer and stomach cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the bone, and the stomach, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., bone, and stomach, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of bone cancer and stomach cancer and possibly other cancers.

FEATURES OF PROTEIN ENCODED BY GENE NO: 61

This gene is expressed primarily in epididymus, and lymph node of breast cancer, and to a lesser extent in many other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, abnormalities of the epididymus, and breast cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the epididymus and breast, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., epididymus and other reproductive tissue, lymphoid tissue, and mammary tissue, and cancerous and wounded tissues) or bodily fluids

15

20

25

30

35

(e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:185 as residues: Arg-57 to Ser-65.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of abnormalities of the epididymus, and breast cancer.

10 FEATURES OF PROTEIN ENCODED BY GENE NO: 62

The translation product of this gene appears to be the human homolog of bovine NADH dehydrogenase which is thought to be important in cellular metabolism. In specific embodiments, the polypeptides of the invention comprise the amino acid sequence: SMSALTRLASFARVGGRLFRSGCARTAGDGGVRHAGGGVHIEPRY RQFPQLTRSQVFQSEFFSGLMWFWILWRFWHDSEEVLGHFPYPDPSQWTDEEL GIPPDDED (SEQ ID NO:323), or fragments thereof. Polynucleotides encoding this polypeptide are also encompassed by the invention.

This gene is expressed in larynx tumor, lymph node, brain amygdala, human cardiomyopathy, and retina.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, diseases affecting cellular metabolism. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., larynx, lymphoid tissue, brain and other tissue of the nervous system, heart and cardiovascular tissue, and retina, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:208 as residues: Pro-27 to Gln-32, Arg-42 to Glu-51.

The tissue distribution and homology to NADH dehydrogenase indicates that

10

15

20

polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of diseases involving cellular metabolism.

FEATURES OF PROTEIN ENCODED BY GENE NO: 63

This gene is expressed primarily in amygdala, and to a lesser extent in many other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, abnormalities of the amygdala. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the amygdala, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., amygdala, and lymphoid tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:187 as residues: Gln-17 to Glu-29, Pro-41 to Phe-46, Ser-59 to Ile-70, Thr-97 to Leu-105.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of abnormalities of amygdala.

25

30

35

FEATURES OF PROTEIN ENCODED BY GENE NO: 64

This gene is expressed primarily in female bladder, and to a lesser extent in chronic synovitis and hemangiopericytoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, bladder cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the urinary tract, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., bladder, synovial tissue, and

vascular tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:188 as residues: Pro-2 to Gln-7, Pro-27 to Phe-34.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatments of defects of the urinary tract, especially bladder cancer.

10

15

20

25

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 65

This gene is expressed primarily in fetal spleen, and to a lesser extent in hemangiopericytoma, thymus, and synovial sarcoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, defects of immune of hematopoietic systems. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune of hematopoietic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., spleen, vascular tissue, thymus, blood cells, and synovial tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The protein product of this gene is useful for treatment of defects of the immune or hematopoietic systems, because of the gene's expression in thymus and spleen.

30

35

FEATURES OF PROTEIN ENCODED BY GENE NO: 66

This gene is expressed primarily in human pituitary and to a lesser extent in placenta and fetal lung.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are

10

15

20

25

30

35

not limited to, endocrine growth disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., pituitary and other endocrine tissue, placental, and pulmonary tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:190 as residues: Val-38 to Asn-44, Gly-53 to Ser-65.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment of growth disorders related to pituitary dysfunction.

FEATURES OF PROTEIN ENCODED BY GENE NO: 67

The translation product of this gene shares sequence homology with a Caenorhabditis elegans gene of unknown function. In specific embodiments, the polypeptides of the invention comprise the sequence: DPRRPNKVLRYKPPPSE CNPALDDPTP (SEQ ID NO:317); DYMNLLGMIFSMCGLMLKLKWCAWVA VYCS (SEQ ID NO:318); FISFANSRSSEDTKQMMSSF (SEQ ID NO:316); and/or MLSISAVVMSYLQNPQPMTPPW (SEQ ID NO:319). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in primary breast cancer and lymph node breast cancer and to a lesser extent in adult brain, lung cancer, colon cancer, epithelioid sarcoma, and Caco-2 cell line.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cancer and tumor tissues, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., mammary tissue, lymphoid tissue, brain and other tissue of the nervous system, lung, colon, and

epithelium, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:191 as residues: Asn-34 to Lys-42.

The tissue distribution in a variety of cancer tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of a variety of cancer and tumor types.

10

15

25

30

35

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 68

The translation product of this gene shares sequence homology with steroid membrane binding protein. The translation product of this gene has recently been published as progesterone binding protein. See Genbank AJ002030. Preferred polypeptides encoded by this gene comprise the following amino acid sequence: AAGDGDVKLGTLGSGSESSNDGGSESPGDAGAAAXGGGWAAAALALLTG GGE (SEQ ID NO:320).

This gene is expressed primarily in breast, and to a lesser extent in placenta and fetal tissue.

20 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, breast cancer or developmental disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of breast or fetal tissues, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., mammary tissue, placenta, and fetal tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:192 as residues: Pro-43 to Asp-49, Gln-54 to Pro-64, Asp-110 to Asp-118, Lys-138 to Tyr-143, Pro-150 to Asp-170.

The tissue distribution and homology to steroid membrane binding protein and to progesterone binding protein indicates that the protein products of this gene are

10

15

20

25

30

35

Brichoold- 1410

useful for treatment of breast cancers, especially those caused by estrogen and progesterone binding.

FEATURES OF PROTEIN ENCODED BY GENE NO: 69

Preferred polypeptides encoded by this gene comprise the following amino acid sequence: AADNYGIPRACRNSARSYGAAWLLLXPAGSSRVEPTQDISISDQLGG QDVPVFRNLSLLVVGVGAVFSLLFHLGTRERRRPHAXEPGEHTPLLAPATAQPL LLWKHWLREXAFYQVGILYMTTRLIVNLSQTYMAMYLTYSLHLPKKFIATIPLV MYLSGFLSSFLMKPINKCIGRN (SEQ ID NO:321).

This gene is expressed primarily in macrophage (GM-CSF treated), and to a lesser extent in monocytes and dendritic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, inflammation and infection. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., macrophages and other blood cells, and dendritic cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that the protein products of this gene are useful for treatment of infection or inflammation or other events or defects involving the immune system.

FEATURES OF PROTEIN ENCODED BY GENE NO: 70

This gene is expressed primarily in adult brain and to a lesser extent in thyroid, 12 week old early stage human, and stromal cell TF274.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, neurological or neuro-endocrine diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes

10

20

25

30

for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous or endocrine systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system, thyroid, and stromal cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:194 as residues: Pro-65 to Cys-71.

The tissue distribution indicates that the protein products of this gene are useful for treatment and diagnosis of neurological diseases or metabolic conditions involving the neuro-endocrine system.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 71

This gene is expressed in T-cell helper and to a lesser extent in adult brain and adult testes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune disorders, meningitis or reproductive problems. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, neural and reproductive systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., T-cells and other blood cells, brain and other tissue of the nervous system, testes and other reproductive tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:195 as residues: Val-18 to Tyr-24, Ala-89 to Asp-99, Asp-104 to Ala-117, Leu-121 to Pro-136.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis immune and

10

15

20

25

30

35

reproductive disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 72

The translated polypeptide of this contig has a high degree of identity with the Ob Receptor-Associated Protein deposited as GenBank Accession No. 2266638. No function has been determined for the Ob Receptor-Associated Protein, however it is expressed upon stimulation of the Ob Receptor by Leptin.

This gene is expressed in T-cells and to a lesser extent in endothelial and bone marrow cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, acute lymphoblastic leukemia, hematapoetic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hematapoetic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., T-cells and other blood cells, endothelial cells, and bone marrow, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:196 as residues: Ser-61 to Trp-70.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of leukemia and other disorders of the primary immune system. In addition, since this gene appears to be related to the Ob Receptor-Related Protein, it is likely that this polypeptide is also involved in the Ob/Leptin signal transduction cascade. As a result, this protein may be of use in the molecular diagnosis and therapeutic intervention of obesity and related disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 73

The translation product of this contig has homology with furin, a protein thought to be a key endopeptidase in the constitutive secretory pathway. The identification and initial characterization of Furin was reported by Takahasi and

10

15

20

colleagues (Biochem Biophys Res Commun 1993 Sep 15;195(2):1019-1026).

This gene is expressed in neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, diseases of the immune system such as allergies, wound healing and antigen recognition. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., neutrophils and other blood cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment of allergies or other immune disorders since neutrophils are an important part of an allergic response. Further, since this protein appears to be related to Furin, it can be used diagnostically and therapeutically to treat secretory protein processing disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 74

This gene is expressed in the frontal cortex.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, of the motor activity and sensory functions that involve the central nervous system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene

expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of neural disorders that affect cognitive functions.

FEATURES OF PROTEIN ENCODED BY GENE NO: 75

The translation product of this gene shares sequence homology with inorganic pyrophophatase which is thought to be important in the catalysis the hydrolysis of diphosphate bonds, chiefly in nucleoside di- and triphosphates and essential enzymes that are important for controlling the cellular levels of inorganic pyrophosphate (PPi). The bovine homolog of this gene has been identified by Yang and Wensel (J. Biol. Chem. 267:24641-24647 (1992)).

This gene is expressed in osteoclastoma cells and to a lesser extent in epithelial cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, osteoporosis and other bone weakening diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skeletal system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., bone, and epithelial cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:199 as residues: Lys-22 to Tyr-28, Asp-64 to Lys-77, Pro-86 to Ile-91, Gln-99 to Pro-119, Tyr-169 to Asp-174, Lys-176 to Gly-181, Trp-189 to Asn-202, Lys-233 to Gly-239, Ser-250 to Asp-257.

The tissue distribution and homology to inorganic pyrophophatase indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of osteoporosis through the removal of bone by demineralization.

5

10

15

20

25

30